Advances in the understanding of Kohn-Sham DFT via the optimised effective potential method

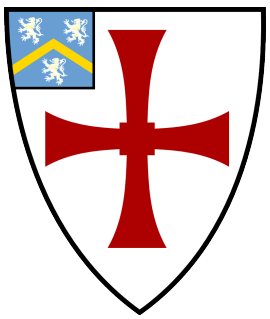
BENJAMIN JAMES PEARCE

Advances in the understanding of Kohn-Sham DFT via the optimised effective potential method

Benjamin James Pearce

A thesis presented for the degree of

Doctor of Philosophy



Department of Physics

The University of Durham

United Kingdom

April 2023

Advances in the understanding of Kohn-Sham DFT via the optimised effective potential method

Benjamin James Pearce

Abstract

Kohn-Sham (KS) density functional theory (DFT) has paved its way to becoming the most widely used method for performing electronic structure calculations. Its major success relies heavily on the underlying approximations that are employed to describe the exchange-correlation (xc) energy functional; hence understanding these approximations proves to be of vital importance.

The main goal of this thesis is to explore and develop a deeper understanding of approximations made within DFT; with a focus on systematically improving existing (semi-)local density functional approximations (DFAs). To do so, we build upon the existing constrained minimisation method, which requires the optimised effective potential (OEP) scheme, improving its implementation and removing one of its major computational bottlenecks.

This thesis also addresses a long-standing question in the field as to why the KS equations of spin-DFT do not reduce to those of DFT in the limit of zero applied magnetic field. A new OEP scheme is derived to construct DFT approximations that yield near spin-DFT accuracy and correct for a systematic error in the exchange energy for open-shell systems. This work is then extended to ensemble systems of varying electron number, where it is shown that (semi-)local approximations can yield non-zero xc derivative discontinuities; an exotic, non-analytic feature of the exact KS potential.

Building on these new OEP formulations, a novel new method for decomposing the molecular screening density into screening densities localised on individual atoms is introduced. This method is shown to yield the predicted but elusive steps in the xc potential as a diatomic dissociates; a very exciting result given that these steps cannot be captured at all from any DFT so far, let alone a (semi-)local DFA.

Declaration

The work in this thesis is based on research carried out in the Condensed Matter Theory Group, Department of Physics, University of Durham, England. No part of this thesis has been submitted elsewhere for any other degree or qualification, and it is the sole work of the author unless referenced to the contrary in the text.

The work presented in this thesis has been published or is due to be submitted in journals – the relevant publications are listed below.

- T.J. Callow, **B.J. Pearce**, T. Pitts, N.N. Lathiotakis, M.J.P. Hodgson and N.I. Gidopoulos. "Improving the exchange and correlation potential in density functional approximations through constraints". *Faraday Discuss.*, **224**, 126-144 (2020).
- T.J. Callow, **B.J. Pearce**, N.N. Lathiotakis, and N.I. Gidopoulos. "Density functionals with spin-density accuracy for open shells". *J. Chem. Phys.*, **156**, 111101 (2022).
- **B.J. Pearce**, and N.I. Gidopoulos. "Approximations for the derivative discontinuity via an ensemble optimised effective potential scheme". *To be submitted* (2023).
- **B.J. Pearce**, and N.I. Gidopoulos. "Localisation of the effective screening density". *To be submitted* (2023).

"*You step into the road, and if you don't keep your feet, there is no knowing where you might be swept off to.*" — J.R.R. Tolkien

Acknowledgements

I would like to first express my sincere gratitude to my supervisor, Nikitas Gidopoulos, who provided me with a great deal of support through a challenging few years. I am extremely thankful for the guidance and patience that he afforded me during my studies. I will certainly miss his fantastic espresso machine, although perhaps not the portable sun lamp, which would consistently burn the back of my head during deep discussions surrounding all things DFT.

I thank my co-supervisor, Stewart Clark, who was always ready to help when I encountered a fiendish programming problem. The notorious whisky-tasting night is certainly one of the most memorable nights of my years at Durham. I also thank Nektarios Lathiotakis for his HIPPO code that I used throughout my studies and for his support in fixing it when I would inevitably break it.

Before the pandemic struck, the office was always a great place to be due to the group of fantastic (and eccentric) people I was lucky enough to work alongside: my thanks to Tim, Tom, Zac, Ben, Matt, and Matjaz for making this a consistent highlight. I would like to say a special thanks Tim for always being in favour of a rogue expedition to Newcastle. I also hope Wine Fridays will continue long after I have left the building.

During my PhD, I was lucky to share a house with Joe, who is solely responsible for my exponential increase in the consumption of Guinness during the lockdown. The uncountable evenings playing video games long into the early hours and the over-extended trip to Japan are my fondest memories of the past 4 years. I also thank all of the friends I have made along the way who made my time in Durham a very enjoyable time indeed.

I sincerely thank my parents, Tim and Briony, my brother Michael, my aunt Pam, and my grandparents, Sheila and Tom: without your support throughout my stud-

ies and the years that led up to it, I would not be where I am today.

Last but not least, I thank my now wife, Charise, who has always been extremely patient and supportive of me. I look forward fondly to the many years ahead of us.

Finally, I thank Durham University and the Engineering and Physical Sciences Research Council (EPSRC) for funding my PhD studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since being pioneered in the fields of nuclear physics and meteorology, numerical simulations have become a vital tool in almost every area of science. From modeling black hole mergers [1], to forecasting the UK economic outlook [2], to predicting the impact of climate change [3], simulations have over the years provided us with a deeper understanding of complex systems and an incredible predictive power. More recently, simulations proved absolutely vital in the prediction of the spread COVID-19 pandemic that swept through communities all around the globe [4]. Being able to develop, test, and validate complex simulations efficiently is therefore extremely important in the modern age.

The work presented in this thesis focuses on the development of intuitive and efficient models within the field of electronic structure theory. This huge area of research studies the behaviour of electronic systems and it is directly responsible for the development of much of the modern technology we use today. In particular, this thesis will be centered around density functional theory (DFT); by far the most popular method used within the electronic structure community. Originally developed back in the 1960s [5, 6], research within DFT has gone on to produce an incredible amount of scientific articles, with more than 100,000 articles published since its inception and the number of articles published per year continuing to increase [7]. It is not however just the sheer number of articles that demonstrates

the importance of DFT research. Two of the ten most cited papers of all time are in fact fundamental works within DFT, and these are the only two physics papers within this top-ten list.

But why is it that DFT has become so popular? The answer is arguably due to three major factors: the speed at which DFT calculations can be performed relative to other methods, the degree of accuracy that can be achieved, and the large amount of time and investment that has been spent developing highly efficient DFT codes. In addition to its vast successes, DFT is not perfect (at least not in practice when approximations are introduced) and does suffer from some issues that the community has found difficult to address. The high availability and low computational power required to run these codes however, has led to perhaps the biggest criticism of DFT - its potentially 'black box' nature. With the overwhelming amount of approximations that have been developed within DFT [8], there is a risk that one can simply utilise the approximation that yields a desired experimental result; hence a fundamental understanding of these approximations is of critical importance.

On the topic of these approximations, it should also be noted that despite the many attempts to oust them, the two most commonly used approximations used today in computational chemistry [9] and in solid state [10] were developed almost 30 years ago. The difficulty in developing approximations that are both more accurate and not overwhelmingly less efficient is likely what led to the explosion of approximations mentioned above. Hence, the work in this thesis aims to take a different approach, by attempting to gain a deeper understanding of the underlying theory of DFT and to improve the commonly used functionals used today in a systematic way.

1.1 Outline of thesis

The work presented in this thesis is structured as follows:

Chapter 2

The underlying problem that must be solved within electronic structure theory is introduced – the many-body problem. Following this, several key concepts such as the Born-Oppenheimer approximation, Hartree-Fock theory, and electron correlation are reviewed.

Chapter 3

This chapter presents a detailed overview of the fundamentals and origin of ground state density functional theory (DFT) via the Hohenberg-Kohn theorems and the Kohn-Sham (KS) scheme. The various classes of density functional approximations developed within DFT are then discussed, as well as how the KS scheme is extended to spin-polarised systems via spin-DFT. An alternative method for calculating the KS potential, known as the optimised effective potential (OEP) scheme, is also introduced and discussed. To conclude the chapter, the implementation of DFT for finite basis sets is reviewed.

Chapter 4

The implementation of the OEP method for finite basis sets is presented. The constrained minimisation method is introduced alongside a more general discussion on how well-behaved OEP solutions can be obtained for finite systems. The implementation of this method was completely rewritten to improve its efficiency and to test whether the positivity constraint enforced in earlier studies is required in practice. Finally, results obtained using the new version of the constrained method are presented and discussed.

Chapter 5

This chapter addresses the current wisdom that approximations in spin-DFT are inherently, or in principle, more accurate than density functional approximations (DFAs) for the description of open-shell systems. A fundamental theoretical question since the introduction of spin-DFT in 1972, is why the spin-DFT KS equations do not reduce to the DFT KS equations when there is zero applied magnetic field. In this chapter, a new set of generalised KS equations are derived which provide a link between DFT and spin-DFT in the limit of zero applied magnetic field. A systematic error in the electron exchange energy that contaminates unpolarised DFAs is identified and a new method for constructing functionals that correct this error is presented. Results obtained using this method are then discussed. Lastly, the validity of efficient approximations that completely bypass the OEP method is explored.

Chapter 6

A new OEP formulation for ensemble systems with fractional average electron number is derived. This new ensemble formulation allows for the calculation of the exchange-correlation derivative discontinuity, a property of the exact KS potential that most (semi-)local DFAs do not possess. It is observed that the differences between ensemble potentials around an integer electron number yield a discontinuous step in the vicinity of the system. It is argued that this step can provide an estimate for the derivative discontinuity.

Chapter 7

A novel method for decomposing the total molecular screening density into screening densities localised on individual atoms is introduced. For small atomic separations, this method is shown to reduce the previous OEP schemes discussed. It is

observed however that for larger separations between atoms, i.e. as systems dissociate, calculated OEP potentials using this method are shifted relative to those calculated using prior OEP formulations; a property of the exact KS potential that is extremely hard to capture with (semi-)local DFAs.

Chapter 8

In this final chapter, general conclusions are presented and ideas for future extensions to the ensemble OEP and localisation of the xc screening hole are discussed.

1.2 Atomic units

To simplify the notation of many of the expressions used throughout this thesis, we adopt Hartree atomic units unless otherwise specified. In this system, the unit of energy is the Hartree (H), where 1 H $\approx$ 27.211 eV. To improve readability, certain results in this thesis are quoted in electron-volts (eV).

This set of units is convenient as all of the following fundamental physical constants ubiquitous to quantum mechanics are set to equal one:

- $\hbar$, reduced Planck's constant;
- m_e, atomic mass unit;
- e, elementary charge;
- a_0, Bohr radius.

Chapter 2

The Many-Body Problem

Over the last century, the study of interacting quantum mechanical systems consisting of electrons and nuclei has become one of the most exciting and important areas of science. These systems are however notoriously difficult to treat, despite knowing the exact equations governing their behaviour. In its most general, time-dependent form, the Schrödinger equation exactly describes such systems. The work presented in this thesis, however, is focused on time-independent systems. Additionally, any relativistic effects are also ignored throughout. In this regime, the Schrödinger equation takes the following well-known form [11]:

$$\hat{H}\Psi = E\Psi, \tag{2.1}$$

where $\hat{H}$ is the time-independent Hamiltonian, Ψ is the time-independent wave-function, and E is the total energy of the system.

A system of N_e electrons and N_n nuclei can be represented by the fully-interacting wave-function:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_\mathrm{e}}; \mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_\mathrm{n}}), \tag{2.2}$$

where $\{\mathbf{x}_i = (\mathbf{r}, \sigma_i)\}$ is the set of spatial and spin coordinates for the electrons and $\{\mathbf{R}_\alpha\}$ is the set of coordinates for the nuclei (we ignore nuclear spin). For this system, the Schrödinger equation can be written as:

$$\hat{H}\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_\mathrm{e}}; \mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_\mathrm{n}}) = E\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_\mathrm{e}}; \mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{N_\mathrm{n}}), \tag{2.3}$$

where the Hamiltonian is given by:

$$\begin{aligned}\hat{H} &= \hat{T}_{\mathrm{n}} + \hat{T}_{\mathrm{e}} + \hat{V}_{\mathrm{nn}} + \hat{V}_{\mathrm{ee}} + \hat{V}_{\mathrm{en}} + \hat{V}_{\mathrm{ext}} \\ &= -\frac{1}{2}\sum_{\alpha}^{N_{\mathrm{n}}} \frac{1}{M_\alpha}\nabla^2_{\mathbf{R}_\alpha} - \frac{1}{2}\sum_{i}^{N_{\mathrm{e}}} \nabla^2_{\mathbf{x}_i} + \frac{1}{2}\sum_{\alpha}^{N_{\mathrm{n}}}\sum_{\beta\neq\alpha}^{N_{\mathrm{n}}} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \\ &\quad + \frac{1}{2}\sum_{i}^{N_{\mathrm{e}}}\sum_{j\neq i}^{N_{\mathrm{e}}} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} - \sum_{\alpha}^{N_{\mathrm{n}}}\sum_{i}^{N_{\mathrm{e}}} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{x}_i|} + \hat{V}_{\mathrm{ext}}, \end{aligned} \tag{2.4}$$

where $\hat{T}_{\mathrm{n}}$ and $\hat{T}_{\mathrm{e}}$ are the kinetic energy terms for the nuclei and electrons, $\hat{V}_{\mathrm{nn}}$ and $\hat{V}_{\mathrm{ee}}$ are the nuclear-nuclear and electron-electron interactions, $\hat{V}_{\mathrm{en}}$ is the electron-nuclear interaction, and $\hat{V}_{\mathrm{ext}}$ arises from any external fields, e.g. an electric field, that may be acting on the system. Only the specific case where external fields are zero and thus $\hat{V}_{\mathrm{ext}} = 0$ however will be considered for the rest of this thesis.

2.1 The Born-Oppenheimer approximation

Finding solutions to the time-independent Schrödinger equation is a massive challenge, and as such, further approximations are made to simplify its solution. One of the most established approximations is the Born-Oppenheimer (BO) approximation which decouples the motion of the electrons from the motion of the nuclei [12]. This treatment of nuclei as classical particles with fixed positions relative to the electrons can be justified by the large relative difference in the masses of the two types of particles. Nuclei are at least three orders of magnitude heavier than an electron and therefore move much slower than the electrons that surround them, so the electrons can be approximated as moving on a potential energy surface generated from the instantaneous positions of the nuclei.

To derive the BO equations, we begin by expanding the total wave-function (2.2) in a series:

$$\Psi(\mathbf{x};\mathbf{R}) = \sum_k \phi_{k,\mathbf{R}}(\mathbf{x})\chi_k(\mathbf{R}), \tag{2.5}$$

where $\{\phi_{k,\mathbf{R}}\}$ are the complete set of electronic wave-functions and $\{\chi_k(\mathbf{R})\}$ are coefficients of expansion that depend on the nuclear coordinates, $\mathbf{R}$. This notation

has also been simplified such that $\mathbf{x} = \mathbf{x}_1, \ldots, \mathbf{x}_{N_\mathrm{e}}$ and $\mathbf{R} = \mathbf{R}_1, \ldots, \mathbf{R}_{N_\mathrm{n}}$.

It is important to note that the electronic wave-functions still depend parametrically on the nuclear coordinates and that these products are not completely separable.

The electronic wave-functions $\{\phi_{k,\mathbf{R}}\}$ are chosen as solutions to the electronic Schrödinger equation:

$$\hat{H}_\mathrm{el}\phi_{k,\mathbf{R}}(\mathbf{x}) = \mathcal{E}_k(\mathbf{R})\phi_{k,\mathbf{R}}(\mathbf{x}), \tag{2.6}$$

where $\hat{H}_\mathrm{el} = \hat{H} - T_\mathrm{n}$ is the electronic Hamiltonian.

Substituting (2.5) into the full Schrödinger Equation 2.3 and then multiplying by $\phi_{k',\mathbf{R}}$ and integrating over the electronic coordinates $\mathbf{x}$ yields a set of coupled equations for the nuclear functions $\chi_k(\mathbf{R})$:

$$\left[\hat{T}_\mathrm{n} + \mathcal{E}_{k'}(\mathbf{R}) - E\right]\chi_{k'}(\mathbf{R}) = \sum_k \Gamma_{k'k}\chi_k(\mathbf{R}), \tag{2.7}$$

where the coupling between different electronic states is given by $\Gamma_{k'k}$. If this coupling to terms $k \neq k'$ is neglected (i.e. keep only the $k = k'$ terms), one obtains the decoupled Born-Oppenheimer nuclear equation:

$$\left[\hat{T}_\mathrm{n} + \mathcal{E}_k(\mathbf{R})\right]\chi_k(\mathbf{R}) = E\chi_k(\mathbf{R}). \tag{2.8}$$

At this point, we can now interpret the function $\chi_k(\mathbf{R})$ as a nuclear wave-function, as $\int \mathrm{d}\mathbf{R}|\chi_k(\mathbf{R})|^2 = 1$.

In most situations the BO approximation is accurate, however, there are certain cases that cause it to break down. Some further understanding of these cases can be gained by studying the form of the coupling terms earlier discarded:

$$\Gamma_{k'k} = \frac{1}{2}\sum_\alpha^{N_\mathrm{n}} \frac{1}{M_\alpha}\left[2\left\langle\phi_{k',\mathbf{R}}\middle|(\nabla_\alpha\phi_{k,\mathbf{R}})\right\rangle\cdot\nabla_\alpha + \left\langle\phi_{k',\mathbf{R}}\middle|\left(\nabla_\alpha^2\phi_{k,\mathbf{R}}\right)\right\rangle\right]. \tag{2.9}$$

The first term in $\Gamma_{k'k}$ is commonly referred to as the first-order non-adiabatic or derivative coupling and the second term is the second-order derivative coupling. The diagonal term ($k = k'$) of the second-order derivative coupling is also commonly

known as the diagonal BO correction (DBOC). Upon first inspection, it may seem that both of these terms are always small due to the presence of the large nuclear masses M_α in the multiplying denominator, although this is not always the case.

Most of the time, the first-order derivative coupling is the dominant term; to understand why it can be rewritten as [13]:

$$\langle \phi_{k',\mathbf{R}} | (\nabla_\alpha \phi_{k,\mathbf{R}}) \rangle = \frac{\langle \phi_{k',\mathbf{R}} | \nabla_\alpha \hat{H}_{\mathrm{el}} | \phi_{k',\mathbf{R}} \rangle}{\mathcal{E}_k(\mathbf{R}) - \mathcal{E}_{k'}(\mathbf{R})}. \tag{2.10}$$

Here it is clear that the first-order derivative coupling term is only small when the electronic energies are well separated. If for certain configurations of $\mathbf{R}$ the energies, $\mathcal{E}_k(\mathbf{R})$ and $\mathcal{E}_{k'}(\mathbf{R})$ are equal, the electronic states are degenerate and the BO approximation breaks down. Gidopoulos and Gross [14] proposed a method for incorporating these effects by constructing non-adiabatic electronic states, however for this thesis no systems are studied that exhibit these features and so the BO approximation can be safely utilised throughout the following chapters.

2.2 Approximating the Wave-function

While we have drastically simplified the problem by decoupling nuclear and electronic motion, the electronic Schrödinger equation is still virtually impossible to solve. As the number of electrons increases, the amount of memory required to represent the electron-electron interaction on a real-space grid increases exponentially; rendering exact calculations for even small atoms off-limits. We must therefore make further approximations to the many-body system to form a problem that can be solved practically.

We now drop the electronic wave-function's explicit parametric dependence on the nuclear coordinates and focus only on the ground state solution. We write the electronic Equation 2.6 simply as:

$$\hat{H}\Psi = \left[\hat{T}_{\mathrm{e}} + \hat{V}_{\mathrm{en}} + \hat{V}_{\mathrm{ee}}\right] \Psi = E\Psi, \tag{2.11}$$

where for conciseness, Ψ will now represent the electronic wave function for the rest of this thesis; rather than $\phi_{k,\mathbf{R}}$ used previously.

Before we proceed let us first specify two constraints that exact solutions to the electronic Schrödinger equation must obey:

1. They must be normalised such that the probability of finding the N_e electrons of the system over all space is exactly equal to one:

$$\int \mathrm{d}\mathbf{x}_1 \int \mathrm{d}\mathbf{x}_2 \cdots \int \mathrm{d}\mathbf{x}_{N_\mathrm{e}} |\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_\mathrm{e}})|^2 = 1. \tag{2.12}$$

2. As electrons are Fermions, the electronic wave-function must be anti-symmetric under the exchange of any pair of spacial and spin electronic coordinates:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_{N_\mathrm{e}}) = -\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_{N_\mathrm{e}}). \tag{2.13}$$

2.2.1 The Hartree Approximation

The most basic approximation to the electronic wave-function that can be made is the Hartree approximation [15–17], which consists of writing the many-body wave-function as a product of single-particle orbitals:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_{N_\mathrm{e}}) = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\ldots\psi_{N_\mathrm{e}}(\mathbf{x}_{N_\mathrm{e}}), \tag{2.14}$$

where each orbital $\psi_i(\mathbf{x}) = \sigma_i(\omega)\phi_i(\mathbf{r})$; a product of spin and spatial orbitals. By definition, the spin orbitals are orthonormal, however, we must enforce this constraint on the spatial orbitals in some manner.

Substituting this approximate wave-function into the expectation value of the electronic Hamiltonian yields the following expression for the total energy:

$$\begin{aligned} E &= \langle\Psi|\, \hat{T}_\mathrm{e} + \hat{V}_\mathrm{en} + \hat{V}_\mathrm{ee}\, |\Psi\rangle \\ &= -\frac{1}{2}\sum_i^{N_\mathrm{e}} \int \mathrm{d}\mathbf{r}\ \phi_i(\mathbf{r})\nabla^2\phi_i(\mathbf{r}) - \sum_\alpha^{N_\mathrm{n}} \int \mathrm{d}\mathbf{r}\ \frac{Z_\alpha \rho(\mathbf{r})}{|\mathbf{R}_\alpha - \mathbf{r}|} \\ &\quad + \frac{1}{2}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\ \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{2}\sum_i^{N_\mathrm{e}} \iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\ \frac{|\phi_i(\mathbf{r})|^2|\phi_i(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}, \end{aligned} \tag{2.15}$$

where the electronic density $\rho(\mathbf{r})$ is defined as:

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_\mathrm{e}} |\phi_i(\mathbf{r})|^2. \tag{2.16}$$

As this form of the total wave-function is separable, each single-particle orbital ϕ_i must satisfy its own coupled single-particle equation, known as the Hartree equation:

$$\left[-\frac{\nabla^2}{2} - \sum_{\alpha}^{N_\mathrm{n}} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}|} + \sum_{j \neq i}^{N_\mathrm{e}} \int \mathrm{d}\mathbf{r}' \, \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \tag{2.17}$$

where the third term on the left-hand side can be separated into two terms:

$$\sum_{j \neq i}^{N_\mathrm{e}} \int \mathrm{d}\mathbf{r}' \, \frac{|\phi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} = \int \mathrm{d}\mathbf{r}' \, \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - \int \mathrm{d}\mathbf{r}' \, \frac{|\phi_i(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}. \tag{2.18}$$

The first of these terms is known as the Hartree potential, commonly written as v_H, and can be thought of as a classical mean-field interaction generated from all the orbitals, including ϕ_i. The second term is known as the self-interaction correction (SIC) term, as it removes the interaction between the i-th electron with itself from the single-particle equations. The eigenvalue for the i-th single-particle orbital ϕ_i is given by ϵ_i.

Although the Hartree equations (2.17) are relatively straightforward to solve and self-interaction free, the approximation for the wave-function does not obey the required anti-symmetry constraint. Therefore, to enforce these strict requirements, more sophisticated approximations for the wave-functions are required.

2.2.2 The Hartree-Fock Approximation

In Hartree-Fock (HF) theory, each electron is treated as an independent particle, with a single-particle wave-function assigned to each electron. The simplest form of the total wave-function that obeys both the normalisation (2.12) and anti-symmetry (2.13) constraints is therefore just an anti-symmetric product of the

single-particle orbitals, known as a Slater determinant [18]:

$$\Psi_{\mathrm{SD}}(\mathbf{x}_1,\ldots,\mathbf{x}_{N_\mathrm{e}}) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \cdots & \psi_1(\mathbf{x}_{N_\mathrm{e}}) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \cdots & \psi_2(\mathbf{x}_{N_\mathrm{e}}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{N_\mathrm{e}}(\mathbf{x}_1) & \psi_{N_\mathrm{e}}(\mathbf{x}_2) & \cdots & \psi_{N_\mathrm{e}}(\mathbf{x}_{N_\mathrm{e}}) \end{vmatrix}. \tag{2.19}$$

An alternative way of writing the Slater determinant is acting with the anti-symmetrising operator $\hat{\mathcal{A}}$ on the simple product of the N_e single-particle orbitals:

$$\Psi_{\mathrm{SD}}(\mathbf{x}_1,\ldots,\mathbf{x}_{N_\mathrm{e}}) = \frac{1}{\sqrt{N!}}\hat{\mathcal{A}}\left[\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\cdots\psi_{N_\mathrm{e}}(\mathbf{x}_{N_\mathrm{e}})\right]. \tag{2.20}$$

When this operator acts on the product of the N_e single-particle orbitals it results in the sum over all possible products of all possible combinations of position coordinates and orbital indices.

Within the Hartree-Fock approximation, the electronic wave-function is represented as a single Slater determinant. The optimal determinant can be found by minimising the total energy of the system over the space of all possible single N_e-electron Slater determinants:

$$E_{\mathrm{HF}} = \min_{\bar{\Psi}\to N_\mathrm{e}} \left\langle \bar{\Psi}\right| \hat{H} \left|\bar{\Psi}\right\rangle, \tag{2.21}$$

where $\bar{\Psi}$ is a normalised anti-symmetric trial Slater determinant and the minimising state is Ψ_{HF}.

This minimisation approach to finding the optimal wave-function is based on the Rayleigh-Ritz variational principle, which states that for any trial wave-function that obeys both the normalisation and anti-symmetry constraints, the expectation value of the Hamiltonian will always be greater than or equal to the ground state energy E_0. Only for the exact wave-function will the equality be satisfied and so we infer that the Hartree-Fock energy, $E_{\mathrm{HF}} = \langle\Psi_{\mathrm{HF}}|\,\hat{H}\,|\Psi_{\mathrm{HF}}\rangle > E_0$, is a strict upper bound to the true ground state energy, for $N_\mathrm{e} > 1$.

To proceed with this minimisation we must first find a representation of the Hartree-Fock energy in terms of the single-particle orbitals. Making use of the Slater-

Condon rules, we can express the Hartree-Fock energy as [18, 19]:

$$\begin{aligned} E_{\mathrm{HF}} &= -\frac{1}{2}\sum_{i}^{N_{\mathrm{e}}}\int \mathrm{d}\mathbf{r}\ \phi_i(\mathbf{r})\nabla^2\phi_i(\mathbf{r}) \\ &\quad - \sum_{\alpha}^{N_{\mathrm{n}}}\int \mathrm{d}\mathbf{r}\ \frac{Z_\alpha \rho(\mathbf{r})}{|\mathbf{R}_\alpha - \mathbf{r}|} \\ &\quad + \frac{1}{2}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\ \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ &\quad - \frac{1}{2}\sum_{i}^{N_{\mathrm{e}}}\sum_{j}^{N_{\mathrm{e}}}\delta_{\sigma_i\sigma_j}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\ \frac{\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})\phi_i(\mathbf{r}')\phi_j^*(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \end{aligned} \tag{2.22}$$

where the electronic density, $\rho(\mathbf{r})$, is defined in Equation 2.16.

The last two terms in the expression are known as the Hartree energy E_{H} and the Fock exchange energy E_{x} respectively, and together represent the electron-electron repulsion. The exchange energy is a consequence of the anti-symmetry of the wave-function, ensuring that the energy remains unchanged under the exchange of any pair of (space-spin) electronic coordinates.

Given this form of the Hartree-Fock energy, we can construct an objective functional to be minimised with respect to variations in the spatial part of the single-particle orbitals that enforces their orthonormality through a set of Lagrange multipliers λ_{ij}:

$$\mathcal{G} = E_{\mathrm{HF}} - \sum_{i}^{N_{\mathrm{e}}}\sum_{j}^{N_{\mathrm{e}}}\lambda_{ij}\int \mathrm{d}\mathbf{r}\ \phi_i^*(\mathbf{r})\phi_j(\mathbf{r}). \tag{2.23}$$

The minimisation of this objective functional yields a set of coupled equations reminiscent of the Hartree equations (2.17):

$$\begin{aligned} &\left[-\frac{\nabla^2}{2} - \sum_{\alpha}^{N_{\mathrm{n}}}\frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}|} + \int \mathrm{d}\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\right]\phi_i(\mathbf{r}) \\ &\qquad - \sum_{j}^{N_{\mathrm{e}}}\delta_{\sigma_i\sigma_j}\int \mathrm{d}\mathbf{r}'\frac{\phi_j^*(\mathbf{r}')\phi_j(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\phi_i(\mathbf{r}') = \sum_{j}^{N_{\mathrm{e}}}\lambda_{ij}\phi_j(\mathbf{r}). \end{aligned} \tag{2.24}$$

We can rewrite these equations as an eigenvalue problem by performing a unitary transformation to the single-particle orbitals:

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{en}}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\right]\phi_i(\mathbf{r}) - \int \mathrm{d}\mathbf{r}'\frac{\rho^{\sigma_i}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\phi_i(\mathbf{r}') = \epsilon_i\phi_i(\mathbf{r}), \tag{2.25}$$

where ϵ_i are the diagonal Lagrange multipliers, $v_{\text{en}}(\mathbf{r})$ is the electron-nuclear potential given by:

$$v_{\text{en}}(\mathbf{r}) = -\sum_{\alpha}^{N_{\text{n}}} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}|}, \tag{2.26}$$

and $\rho^{\sigma_i}(\mathbf{r}, \mathbf{r}')$ is the one-body density matrix for spin-σ_i:

$$\rho^{\sigma_i}(\mathbf{r}, \mathbf{r}') = \sum_{j}^{N_{\text{e}}} \delta_{\sigma_i \sigma_j} \phi_j^*(\mathbf{r}) \phi_j(\mathbf{r}'). \tag{2.27}$$

These single-particle equations are known as the Hartree-Fock equations [20]. They can be written more compactly by defining the non-local Fock operator $\hat{F}$ such that:

$$\hat{F}\phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \tag{2.28}$$

Like most sets of simultaneous equations in electronic structure theory, the Hartree-Fock equations must be solved self-consistently via an iterative procedure, known as the self-consistent field (SCF) procedure. This is because the Fock operator is itself constructed from the single-particle orbitals via the density and the density matrix. The general outline of this process is as follows [21]:

1. Expand the Hartree-Fock orbitals in a basis set: $\phi_i(\mathbf{r}) = \sum_k c_{ik}\theta_k(\mathbf{r})$ and define the matrix constructed from the set of coefficients $\{c_{ik}\}$ as $\mathbf{C}$.

2. Construct the overlap matrix $\mathbf{S}$ with elements $S_{kl} = \langle \theta_k | \theta_l \rangle$.

3. Construct the Fock matrix $\mathbf{F}$ with elements $F_{kl} = \langle \theta_k | \hat{F} | \theta_l \rangle$.

4. Solve the Roothaan-Hall equations $\sum_l F_{kl} c_{li} = \epsilon_i \sum_l S_{kl} c_{li}$ for the coefficients c_{ik} and orbital energies ϵ_i (or in matrix form $\mathbf{FC} = \mathbf{SC}\epsilon$, where ϵ is a diagonal matrix constructed from the set $\{\epsilon_i\}$) [22, 23].

5. Repeat steps 2-4 until you reach convergence in both the electronic density and the total energy to within a certain tolerance.

2.2.2.1 Properties of Hartree-Fock theory

Despite providing an improvement to the Hartree approximation by correctly incorporating the anti-symmetric nature of the wave-function, the Hartree-Fock equations are far more computationally demanding to solve. This is due to the non-local Fock exchange term which causes Hartree-Fock to scale in complexity as N_e^4, unlike the Hartree approximation with the SIC term omitted which scales as N_e^3.

It is also important to note that the Hartree-Fock approximation treats the occupied and unoccupied orbitals rather differently. For the lowest energy N_e orbitals, self-interaction terms are exactly cancelled out and so each electron is correctly repelled by the other $N_\mathrm{e} - 1$ electrons. However, each of the unoccupied/virtual orbitals feels a repulsion from all the electrons and therefore experiences a repulsion by N_e electrons. Therefore, we can instead interpret the unoccupied HF orbitals as orbitals of an $N_\mathrm{e} + 1$ electron system. A direct consequence of this is that the unoccupied orbital eigenvalues obtained from Hartree-Fock calculations provide a very poor estimation of excitation energies since they are supposed to give electron affinities (although not particularly accurately).

The Hartree-Fock approximation does however provide a good estimate for the ionisation energies of a system [24]. If we assume that upon the removal of an electron from the i-th orbital of a system all the orbitals are fixed or 'frozen', then Koopmans' theorem states that the total energy of the system after an electron is removed is given by $E^{N_\mathrm{e}-1} = E^{N_\mathrm{e}} - \epsilon_i$ [24]. Combined with the definition of the ionisation energy, we can then say that:

$$I_i = E^{N_\mathrm{e}-1} - E^{N_\mathrm{e}} = -\epsilon_i. \tag{2.29}$$

Ionisation energies obtained in this manner tend to overestimate the exact values obtained via experiments. This is caused by two competing errors of slightly different magnitudes:

1. A lack of any orbital relaxation effects due to invoking the 'frozen' orbital

approximation.

2. A lack of any correlation terms in the underlying Hartree-Fock theory.

Relaxation effects are fairly easy to understand. Within the 'frozen' orbital approximation we are assuming that the orbitals of the resulting $N_\mathrm{e} - 1$ system after the removal of an electron are identical to those of the N_e system before the electron was removed. This approximate energy will always be greater than the relaxed HF ground state energy of the $N_\mathrm{e} - 1$ system due to the variational principle, therefore the resulting energy difference will be smaller. Clearly, this will then result in a lower ionisation energy. The magnitude of these effects varies depending on the number of electrons in the system, and the 'depth' of the electron within the occupied states.

The other error is much harder to define exactly, and it is responsible for many failures of commonly used methods within electronic structure theory. Let us first discuss what we mean by correlation terms. Within Hartree-Fock theory, we model the electronic wave-function as a single Slater determinant. The problem in this approach is that a single Slater determinant cannot represent a fully-interacting electron system. In fact, this form of the wave-function can only exactly represent a non-interacting system, meaning that even if we obtain the optimal determinant we can never recover the exact total energy of the interacting system. The missing energy that accounts for the interacting nature of the electrons is called the correlation energy. This energy is usually much smaller than other energy contributions, such as the Hartree or exchange energies, normally accounting for $\approx$1% of the total energy of the system.

If we study the form of the Hartree-Fock equations (2.25), we can see that each electron effectively 'sees' a mean-field potential that has no explicit dependence on the exact positions of all the other electrons. In a true interacting system, the potential that an electron experiences should take into account these exact positions, which will result in a diminished probability of finding an electron in

the neighbourhood of another electron. This type of correlation error is known as 'dynamic' correlation. As Hartree-Fock does not contain any description of this effect (except by anti-symmetry for same-spin electrons), the electrons are not directly biased by the exact positions of the others and often over-localise in areas where they should not. As the number of electrons in the system increases, the magnitude of this effect also increases, meaning it will have a larger effect on the energy of the N_e system than on the $N_e - 1$ system. As this effect increases the magnitude of total energies, this error will work to increase the magnitude of calculated ionisation energies; in contrast to relaxation to effects that do the opposite.

For the unoccupied orbitals, which can be thought of as electron affinities, these two errors however do not almost cancel out, but instead add up, resulting in less accurate predictions for electron affinities relative to ionisation energies.

2.2.2.2 Restrictions on electron spin

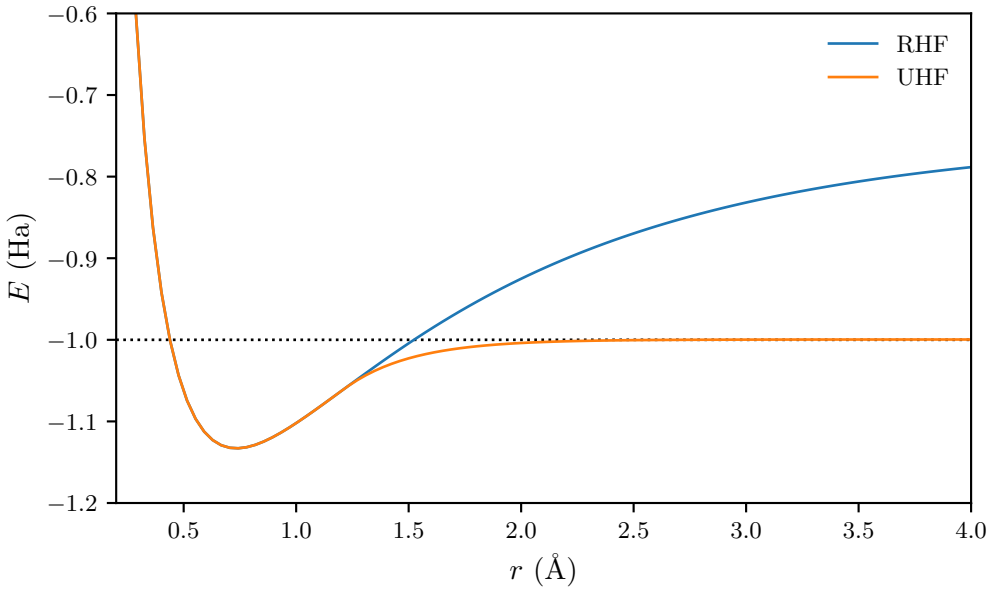


Figure 2.1: Comparison of the molecular dissociation curves for H_2 obtained from RHF and UHF. The point where the two solutions deviate is known as the Coulson-Fischer point [25].

To conclude our discussion of Hartree-Fock theory, we consider a further restriction that can be placed on the Hartree-Fock orbitals if the system in question is closed-shell, i.e. has an equal number of spin-up and spin-down electrons. For such systems, one can make the assumption that each of the orbitals are doubly occupied containing electrons of opposite spins. This approach is called restricted Hartree-Fock (RHF). If this restriction is not applied and electrons of opposite spins do not necessarily occupy the same orbital, the method is known as unrestricted Hartree-Fock (UHF), which lends itself well to open-shell calculations, often with an odd number of electrons.

In the case of open-shell systems, one could also apply the same doubly occupied restriction to all of the orbitals except the one that contains the highest energy electron that is unpaired. As this method is essentially an extension of RHF for an open-shell case it is called restricted open-shell Hartree-Fock (ROHF) [26].

For closed-shell systems the difference in the total energies calculated from RHF or UHF is small, however, the wave-function obtained via UHF is not, in general, an eigenfunction of the total spin-squared operator $\hat{S}^2$ and hence contains some amount of artificial mixing between different electronic spin states. This wave-function, therefore, is said to suffer from spin-contamination, which can cause several issues in calculations [27, 28]. As such, when performing Hartree-Fock calculations for a closed-shell system RHF is generally more commonly used. However, for certain applications, such as when we stretch a bond in a closed-shell system, the spin-contamination found in the UHF wave-function can be a benefit, as it is able to partially capture a specific case of a 'static' correlation effect by breaking the symmetry of the solution. A classic historical example is the dissociation curve of a H_2 molecule. In RHF, as we stretch the molecule there is an equal probability of the dissociation yielding two ions or two atoms. Physically, the diatomic should split into two atoms, not two ions. As UHF is able to break spin symmetry, it allows the molecule to dissociate into one hydrogen atom with a spin-up electron, and one with a spin-down electron. This can be seen in Figure 2.1, where UHF is able to

correctly model the H_2 molecule's accurate energetic behaviour in its dissociation limit.

2.2.3 Post-HF methods

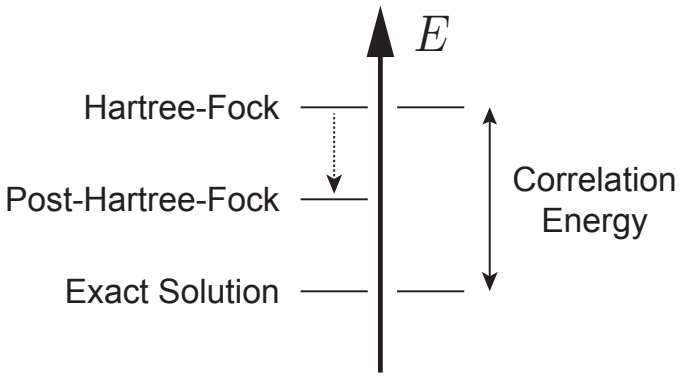


Figure 2.2: Definition of electron correlation energy in terms of the energy difference between the Hartree-Fock energy and the exact interacting energy.

During our discussion of Hartree-Fock theory, we highlighted that approximating the wave-function as a single Slater determinant consisting of products of single-particle orbitals fails to include the effect of electron correlation. By definition, Hartree-Fock models electron exchange exactly and so any method aiming to improve upon HF essentially needs to be able to accurately approximate the interacting nature of the electrons in a system. Methods that improve upon Hartree-Fock methods are collectively known as post-Hartree-Fock methods.

In Figure 2.2, we can see how electron correlation energy is defined in terms of the energy difference between the Hartree-Fock energy and the exact interacting energy. Post-Hartree-Fock methods aim to capture some of the difficult correlation effects and in general provide more accurate results. This improvement comes with a cost, however, as this collection of methods are computationally more demanding than regular Hartree-Fock calculations.

Four of the most commonly used post-Hartree-Fock methods are:

- Configuration interaction (CI) methods [29].
- Møller–Plesset (MP) perturbation theory [30].
- Coupled cluster (CC) theory [31].
- Stochastic methods, e.g. Quantum Monte Carlo [32]

All four of these methods are based on an expansion of the electronic wave-function in terms of multiple Slater determinants corresponding to different excited states, but their ansatz for the wave-function takes different forms depending on the method in question. As expected, as the amount of excitation levels included in these methods increases the computational cost increases sharply, meaning that the application of these correlated methods is limited to systems with a small number of electrons.

Chapter 3

Ground State Density Functional Theory

Having spent a great deal of time discussing how we can approximate the electronic wave-function to achieve accurate solutions to the Schrödinger equation it is time to revert to the beginning of the previous chapter and formulate a radically different approach to modelling electronic systems. Rather than focusing on finding an accurate approximation to the wave-function, density functional theory (DFT) makes use of electron density; which has many practical and theoretical advantages that are discussed in this chapter.

Here the focus is specifically on DFT for obtaining the ground state energy of a system, presenting a historical overview of the development of the most commonly used approach for solving the electronic structure for molecular and solid-state systems.

3.1 The Thomas-Fermi model

Before the development of Hartree-Fock theory, another method was developed with the aim of providing a rough approximate solution to the electronic structure in terms of the electronic density, known as the Thomas-Fermi method [33, 34].

Rather than focusing on approximating the wave-function, Thomas and Fermi postulated that the electronic density could be thought of as the fundamental variable. Unlike the complexity of the wave-function which scales exponentially with electron number, the electronic density scales linearly, and so if we can construct a theory that does not require the construction of a wave-function a huge computational advantage is achieved.

Thomas and Fermi proposed that the total energy of an electronic system could be written as a functional of the electronic density composed of three terms:

$$E_{\mathrm{TF}}[\rho(\mathbf{r})] = A_k \int \mathrm{d}\mathbf{r}\, [\rho(\mathbf{r})]^{\frac{5}{3}} + \int \mathrm{d}\mathbf{r} \rho(\mathbf{r}) v_{\mathrm{en}}(\mathbf{r}) + \frac{1}{2} \iint \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{3.1}$$

where the first term is an approximation of the kinetic energy modelled on the non-interacting homogeneous electron gas (HEG); the second is the classical electron-nuclei interaction energy, and the third is the Hartree energy representing the electron-electron repulsion energy.

Using again the method of Lagrange multipliers to minimise the Thomas-Fermi energy under the constraint that the electronic density integrates to the number of electrons N_{e} (with a multiplier of μ), yields a single equation that can be solved to obtain the electronic density:

$$\frac{5}{3} A_k \, [\rho(\mathbf{r})]^{\frac{2}{3}} + v_{\mathrm{en}}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - \mu = 0. \tag{3.2}$$

The Thomas-Fermi form of the total energy (3.1) has several serious problems, including the absence of terms representing electron exchange or correlation. Its most severe shortcoming however is its crude approximation to the kinetic energy functional [35]; this is because building the fact that electrons must satisfy the Pauli exclusion principle and occupy single-particle states one-by-one into any expression that depends only on the total electronic density is extremely difficult. As the kinetic energy of an electronic system contributes a significant percentage of the total energy, any errors in its approximation can lead to much larger errors than say by neglecting correlation effects. Despite having no real-world usefulness,

the Thomas-Fermi method was the start of deep discussions within the electronic structure theory community around the role of the electronic density. Several decades after Thomas and Fermi published their method, the breakthrough work by Hohenberg and Kohn [5] proved that the formulation of an exact theory, based on the electronic density, is actually possible.

3.2 Hohenberg-Kohn theorems

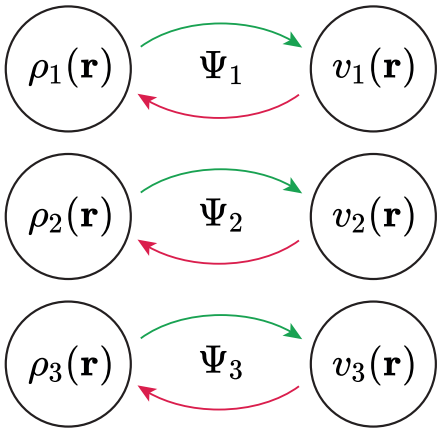


Figure 3.1: A schematic representation of the one-to-one mapping between electronic densities and external potentials established by the Hohenberg-Kohn theorems.

Although it is clearly computationally favourable to use the electronic density as the key variable when attempting to solve the Schrödinger equation, is it actually justified in doing so? The solution to this fundamental question was provided by the Hohenberg-Kohn (HK) theorems [5]. Together these two theorems paved the way for the development of the most popular and significant approaches to solving electronic structure problems. The two HK theorems and their proofs are as follows:

1. The electron-nuclear potential $v_{\text{en}}(\mathbf{r})$ is a unique functional of the total electronic density $\rho(\mathbf{r})$.

To prove this theorem we must first show that no two potentials that differ by more than a simple constant c can yield the same ground state wave-function (eigenfunction in general).

Let's assume that we have two different systems, each with an external potential $v(\mathbf{r})$ and $v'(\mathbf{r})$ respectively, that differ by more than a constant and share the same wave-function Ψ_v. These two systems satisfy their own Schrödinger equations:

$$\hat{H}_v \Psi_v = \left[\hat{T} + \hat{V}_{\text{ee}} + \hat{v}\right] \Psi_v = E_v \Psi_v, \tag{3.3}$$

$$\hat{H}_{v'} \Psi_v = \left[\hat{T} + \hat{V}_{\text{ee}} + \hat{v}'\right] \Psi_v = E_{v'} \Psi_v ., \tag{3.4}$$

where from this point on the electronic kinetic energy operator will simply be written as $\hat{T}$.

By subtracting Equation 3.4 from Equation 3.3 we find:

$$\left[\hat{v} - \hat{v}'\right] \Psi_v = \left[E_v - E_{v'}\right] \Psi_v = c\Psi_v, \tag{3.5}$$

where $c = E_v - E_{v'}$.

As both of the potentials are multiplicative, if we assume that the wave-function $\Psi_v \neq 0$ we can divide the above equation by Ψ_v:

$$\hat{v} - \hat{v}' = \sum_k \left[v(\mathbf{r}_k) - v'(\mathbf{r}_k)\right] = c. \tag{3.6}$$

Clearly, this is a contradiction to our initial assumption, that the potentials differ by more than a constant. This completes the first half of our proof.

The next step is to prove that each ground-state wave-function arising from a potential yields a unique ground-state electronic density. Using the Rayleigh-Ritz variational principle, we know that for two systems with different wave-functions Ψ_v and $\Psi_{v'}$ the following inequalities must be satisfied:

$$\langle \Psi_v | \hat{H}_v | \Psi_v \rangle < \langle \Psi_{v'} | \hat{H}_v | \Psi_{v'} \rangle , \tag{3.7}$$

$$\langle \Psi_{v'} | \hat{H}_{v'} | \Psi_{v'} \rangle < \langle \Psi_v | \hat{H}_{v'} | \Psi_v \rangle . \tag{3.8}$$

We can then take the difference between these two inequalities to find:

$$\int \mathrm{d}\mathbf{r} \left[v'(\mathbf{r}) - v(\mathbf{r})\right] \left[\rho_{v'}(\mathbf{r}) - \rho_v(\mathbf{r})\right] < 0, \tag{3.9}$$

where $\rho_v(\mathbf{r})$ is the ground-state density of H_v.

By setting $\rho_v(\mathbf{r}) = \rho_{v'}(\mathbf{r})$ the inequality above reads $0 < 0$, which is clearly absurd and thus completes the second part of this proof (where we have assumed no degeneracy).

Having established the one-to-one mapping between the external potential and the electronic density, the total energy can be expressed as an explicit functional of the electronic density. By defining the Hohenberg-Kohn energy functional $F_{\mathrm{HK}}[\rho] = \langle \Psi_\rho | \, \hat{T} + \hat{V}_{\mathrm{ee}} \, | \Psi_\rho \rangle$, the total energy for a given external potential is:

$$E_v[\rho] = F_{\mathrm{HK}}[\rho] + \int \mathrm{d}\mathbf{r} \; v(\mathbf{r})\rho(\mathbf{r}) = \langle \Psi_\rho | \, \hat{T} + \hat{V}_{\mathrm{ee}} + v \, | \Psi_\rho \rangle \, . \tag{3.10}$$

From this expression, we deduce that knowledge of the exact form of the functional $F_{\mathrm{HK}}[\rho]$ would allow one to obtain the exact ground state density and energy of the electronic system. Sadly, knowledge of the exact form of $F_{\mathrm{HK}}[\rho]$ is not possible; as we would need to effectively solve a problem as complex as the one we are attempting to avoid. Despite this, several major advances have been made in allowing the form of $F_{\mathrm{HK}}[\rho]$ to be approximated to practical degrees of accuracy, which are discussed later in this chapter.

2. The exact ground state density $\rho_0(\mathbf{r})$ is the density that minimises the total energy functional $E_v[\rho]$.

   This proof can be completed easily by making use of the variational principle. Let's assume that we have an N electron system in $v(\mathbf{r})$ with ground state wave-function Ψ_0 and ground state density $\rho_0(\mathbf{r})$. The ground state energy E_0 of such a system is given by:

$$E_0 = \langle \Psi_0 | \, \hat{H}_v \, | \Psi_0 \rangle = \langle \Psi_0 | \, \hat{F}_{\mathrm{HK}} + \hat{v} \, | \Psi_0 \rangle \, . \tag{3.11}$$

By the variational principle, we know that for any density $\rho(\mathbf{r})$ arising from an external potential $v(\mathbf{r})$ and wave-function $\Psi \neq \Psi_0$, the value of the total energy functional $E_v[\rho]$ is greater than E_0. Only if $\Psi = \Psi_0$ does the value of $E_v[\rho] = E_0$, i.e. $E_v[\rho] \geq E_v[\rho_0]$. Therefore, the ground state density can be determined by searching for the density that minimises the total energy functional. In practice, this can be achieved by searching over all normalised N electron wave-functions for the one that minimises the expectation value of $\hat{H}_v$:

$$E_0 = \min_{\Psi \to N} \langle \Psi | \hat{H}_v | \Psi \rangle . \tag{3.12}$$

3.3 Constrained search formulation

Despite their success in cementing the electronic density as the fundamental variable, it is important to note a shortcoming of the above Hohenberg-Kohn proofs, that they are only defined for densities that are v-representable, i.e. those that are ground states of a Hamiltonian with a continuous local potential $v(r)$. Both of these issues can be alleviated by considering the constrained search formulation of Levy and Lieb [36–38].

The constrained search approach bypasses the v-representability problem of the regular HK proofs by instead enforcing the less strict condition that the density must be N-representable [39], i.e. there exists an N-electron anti-symmetric wave-function with that density. The main idea of this approach is to separate the minimisation (3.12) into two separate minimisations. The first of these involves minimising the expectation value of $\hat{H}_v$ over all N electron wave-functions that yield a particular density ρ:

$$E_v[\rho] = \min_{\Psi \to \rho} \langle \Psi | \hat{H}_v | \Psi \rangle . \tag{3.13}$$

Next, we minimise the functional $E_v[\rho]$ by searching over all possible densities that

integrate to N electrons:

$$E_0 = \min_{\rho \to N} E_v[\rho] = \min_{\rho \to N} \left\{ \hat{F}[\rho] + \int \mathrm{d}\mathbf{r}\ v(\mathbf{r})\rho(\mathbf{r}) \right\}, \tag{3.14}$$

where the universal functional $F[\rho]$ is given by:

$$F[\rho] = \min_{\Psi \to \rho} \langle \Psi | \, \hat{T} + \hat{V}_{\mathrm{ee}} \, | \Psi \rangle \, . \tag{3.15}$$

The minimum of $E_v[\rho]$ is only obtained when $\rho = \rho_0$, so when this condition is met we know that the functional derivative under the constraint that the density must integrate to the electron number N must be zero, leading us to the DFT Euler equation:

$$v_{\mathrm{en}}(\mathbf{r}) = -\frac{\delta F[\rho]}{\delta \rho(\mathbf{r})}\bigg|_{\rho=\rho_0} + \mu. \tag{3.16}$$

The presence of the Lagrange multiplier μ enforces the constraint that the density must integrate to N and so we can write that $\mu = \mu(N)$. By extending ground state DFT to ensemble DFT (with states having different numbers of electrons), μ can be identified as the chemical potential of the system, given by the partial derivative of E_0 with respect to the number of electrons [40]:

$$\mu = \frac{\partial E_0}{\partial N}. \tag{3.17}$$

3.4 Kohn-Sham theory

The real stumbling block in pursuing DFT up to this point is, like Thomas-Fermi theory, finding an accurate approximation for the kinetic energy functional. The Kohn-Sham (KS) scheme cleverly decides this is a battle worth conceding and instead opts to avoid this problem altogether [6]. Using the ideas put forward in the HK theorems, Kohn and Sham mapped the interacting system to an auxiliary system of non-interacting electrons, represented by a single Slater determinant, which has the exact ground state density as the true interacting system. This auxiliary system is commonly known as the KS system. To enforce the same densities, the

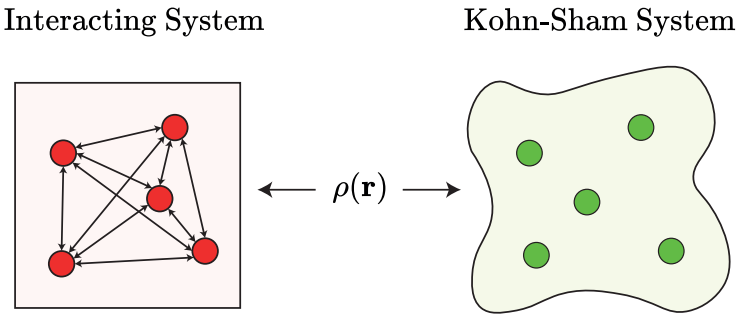


Figure 3.2: A graphical illustration of the Kohn-Sham scheme, showing how the interacting system is mapped to a non-interacting system that retains the same electronic density by modifying its potential.

KS system experiences a different potential than the interacting system, called the KS potential $v_{\mathrm{s}}(\mathbf{r})$ (see Figure 3.2 for a pictorial view of these ideas).

At first, this approach may seem counter-intuitive, as we appear to be moving further away from finding the exact solution. However, there is a major advantage to this approach, which lies in the knowledge of how to construct the kinetic energy functional for a single Slater determinant Φ. From HF theory, the kinetic energy functional is given by:

$$T_{\mathrm{s}}[\rho] = \langle \Phi | \hat{T} | \Phi \rangle = -\frac{1}{2} \sum_{i=1}^{N} \int \mathrm{d}\mathbf{r} \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}). \tag{3.18}$$

Although this kinetic energy is not exactly the same as that of the interacting system, it provides a far better approximation than the Thomas-Fermi model or other orbital-free methods. We note here that the kinetic energy is not an explicit functional of the density, in that the density does not appear directly in the definition of the functional. Instead, its dependence on the density is implicit due to its explicit dependence on the KS orbitals. Such functionals are known as orbital-dependent or implicit density functionals, and they form the basis for the majority of the work later in this thesis.

As we did for $F[\rho]$ in the previous section, the kinetic energy functional can be defined for this auxiliary non-interacting system as the minimisation over all pos-

sible Slater determinants that yield a density ρ:

$$T_{\mathrm{s}}[\rho] = \min_{\Phi \to \rho} \langle \Phi | \, \hat{T} \, | \Phi \rangle \, . \tag{3.19}$$

The single-particle orbitals of the KS system that experience a potential $v_{\mathrm{s}}[\rho](\mathbf{r})$ must obey their own set of single-particle Schrödinger equations, known as the KS equations:

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{s}}[\rho](\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \tag{3.20}$$

Following the same method as for the interacting system, an Euler equation for the KS system can be derived by minimising the kinetic energy functional under the constraint that $\int \mathrm{d}\mathbf{r} \rho(\mathbf{r}) = N$:

$$v_{\mathrm{s}}[\rho](\mathbf{r}) = -\frac{\delta T_{\mathrm{s}}[\rho]}{\delta \rho(\mathbf{r})} + \mu_{\mathrm{s}}. \tag{3.21}$$

We must now work out how to directly relate the KS potential to the fully interacting system. To account for the difference between the interacting and non-interacting kinetic energy functionals, an additional remainder term must be introduced called the exchange-correlation functional:

$$E_{\mathrm{xc}}[\rho] = F[\rho] - T_{\mathrm{s}}[\rho] - E_{\mathrm{H}}[\rho], \tag{3.22}$$

where the Hartree energy functional is given by:

$$E_{\mathrm{H}}[\rho] = \frac{1}{2} \iint \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{3.23}$$

Together with $T_{\mathrm{s}}[\rho]$, and the expressions for the electron-nuclear and Hartree energies, the total energy functional of the interacting system can be rewritten as:

$$E_v[\rho] = T_{\mathrm{s}}[\rho] + \int \mathrm{d}\mathbf{r} \; v_{\mathrm{en}}(\mathbf{r}) \rho(\mathbf{r}) + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}[\rho]. \tag{3.24}$$

In this form, the exchange-correlation density functional contains the only remaining unknown information of the universal functional. It accounts not only for the correction to the kinetic energy but also for electron exchange (which is included exactly in HF) and correlation effects.

Substituting the definition of $E_{\mathrm{xc}}[\rho]$ into the DFT Euler Equation 3.16 we find:

$$v_{\mathrm{en}}(\mathbf{r}) = -\frac{\delta T_{\mathrm{s}}[\rho]}{\delta\rho(\mathbf{r})} - v_{\mathrm{H}}[\rho](\mathbf{r}) - v_{\mathrm{xc}}[\rho](\mathbf{r}) + \mu, \tag{3.25}$$

where the exchange-correlation potential is given by the functional derivative:

$$v_{\mathrm{xc}}[\rho](\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r})}. \tag{3.26}$$

We can then rearrange Equation 3.21 and substitute it into Equation 3.16, to re-express the KS potential as:

$$v_{\mathrm{s}}[\rho](\mathbf{r}) = v_{\mathrm{en}}(\mathbf{r}) + v_{\mathrm{H}}[\rho](\mathbf{r}) + v_{\mathrm{xc}}[\rho](\mathbf{r}) + c, \tag{3.27}$$

where $c = \mu_{\mathrm{s}} - \mu$. The Lagrange multiplier μ_{s} acts similarly to μ for the interacting system, in that it enforces the constraint that the density must integrate to N. It is clear from this equation therefore that in DFT, the KS potential is determined with the freedom of a constant.

We now have an equation that allows us to directly relate the KS potential to the fully interacting system. With both the KS equations (3.20) and Equation 3.27, we have in principle everything required to determine the exact ground state density of the fully interacting system. As with other methods previously discussed, the KS equations must be solved self-consistently as constructing the KS potential requires knowledge of the KS orbitals and vice versa. We explore the implementation of these equations in numerical calculations later in this chapter.

However, one unknown remains, the exchange-correlation functional. Having dramatically reduced the error in the kinetic energy functional, we must still take care in how we approximate $E_{\mathrm{xc}}[\rho]$. The following section presents a historical overview of the important stages in the development of accurate exchange-correlation functionals.

3.5 Exchange-correlation functionals

Within the KS scheme of DFT, the exchange-correlation functional holds all of the remaining unknown information that prevents us from finding exact solutions for the ground-state density and energy of the electronic problem. Therefore, it is imperative that accurate approximations of this functional must be obtained. In this section, we introduce the major classes of approximations to $E_{\mathrm{xc}}[\rho]$, and how these classes can be roughly arranged in order of accuracy and computational cost.

3.5.1 Local density approximation

Local density approximations are an important type of density functional approximation where the exchange-correlation energy at any point in space is only dependent on the electronic density at that point [5]. There are a variety of different approaches that can be used to develop LDAs, however, most of these are based on the homogeneous electron gas (HEG). We can write a LDA in the following way:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho] = \int \mathrm{d}\mathbf{r}\ \epsilon_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})]\rho(\mathbf{r}), \tag{3.28}$$

where $\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})]$ is the exchange-correlation energy-density, which can be decomposed into the exchange and correlation energy-densities:

$$\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})] = \epsilon_{\mathrm{x}}^{\mathrm{LDA}}[\rho(\mathbf{r})] + \epsilon_{\mathrm{c}}^{\mathrm{LDA}}[\rho(\mathbf{r})]. \tag{3.29}$$

Within the LDA, we can define the exchange-correlation potential simply in terms of the exchange-correlation energy-density by constructing the functional derivative:

$$v_{\mathrm{xc}}^{\mathrm{LDA}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}^{\mathrm{LDA}}}{\delta\rho(\mathbf{r})} = \epsilon_{\mathrm{xc}}^{\mathrm{LDA}}(\rho) + \rho(\mathbf{r})\frac{\partial\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}. \tag{3.30}$$

As the exchange energy-density can be analytically determined for the HEG (see the HF solution for jellium in Ref. [41]), we can write the exchange energy-density as:

$$\epsilon_{\mathrm{x}}^{\mathrm{LDA}}[\rho](\mathbf{r}) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\rho(\mathbf{r}). \tag{3.31}$$

Finding an accurate expression for the correlation energy-density is slightly more challenging. In the HEG model, $\epsilon_{\mathrm{c}}^{\mathrm{LDA}}[\rho(\mathbf{r})]$ takes two different forms in the limits of high and low density, which in turn correspond to the limits of weak and strong correlation respectively. In the high-density limit, it can be shown that $\epsilon_{\mathrm{c}}[\rho(\mathbf{r})]$ can be written as:

$$\epsilon_{\mathrm{c}}^{\mathrm{LDA}}[\rho(\mathbf{r})] = c_1 \ln(r_{\mathrm{s}}) + c_2 + r_{\mathrm{s}} \left[c_3 \ln(r_{\mathrm{s}}) + c_4\right], \tag{3.32}$$

where the density dependence enters via the Winger-Seitz parameter r_{s}, defined as the radius of a sphere containing exactly one electron, divided by the Bohr radius a_0. In terms of the electronic density, it can be expressed as:

$$r_{\mathrm{s}} = \left(\frac{3}{4\pi\rho}\right)^{\frac{1}{3}}. \tag{3.33}$$

Correspondingly, in the low-density limit the correlation energy-density is given by:

$$\epsilon_{\mathrm{c}}^{\mathrm{LDA}}[\rho(\mathbf{r})] = \frac{1}{2}\left[\frac{g_0}{r_{\mathrm{s}}} + \frac{g_1}{r_{\mathrm{s}}^{3/2}} + \ldots\right]. \tag{3.34}$$

By interpolating between these two limiting expressions and comparing against the results of quantum Monte-Carlo calculations, several different parameterisations of the correlation energy-density have been constructed, which each yield subtly different numerical results [42–44].

LDAs based on the HEG may seem very reminiscent of the crude Thomas-Fermi approximation, which yields poor results. However, in these approximations, only the exchange-correlation energy is estimated in this way, not the much larger kinetic energy. The LDA yields good results for systems where the electronic density is near-uniform, i.e. where the electrons are delocalised over large regions of space. For solid-state calculations on metallic systems, the LDA tends to work especially well. When considering finite molecular systems, however, where the electronic cloud is more strongly localised to the atoms, the LDA does not perform so well, with a severe tendency to overbind molecules and the prediction of unbound ions that are stable. One exception is that calculated total energies are usually within

a few percent of the experimental values. The major benefit of the LDA is that it is extremely computationally efficient, especially when compared to more complex approximations we will introduce shortly.

3.5.2 Generalised gradient approximations

An approach to constructing more sophisticated approximations for exchange-correlation density functionals was in fact proposed by Hohenberg and Kohn in their original paper [5]. They introduced an extension to the LDA which they termed the gradient expansion approximation (GEA). Although this approach initially provided very disappointing results due to the failure to satisfy several conditions on the exchange-correlation hole (the space around each electron that reduces the probability of finding other electrons in that neighbourhood). Work by Perdew et al. provided a way to fix the issues present in the GEA [45], which led to the construction of arguably the most successful (to date) class of functionals known as the generalised gradient approximations (GGAs) [46–48]. In general, the exchange-correlation functional for a GGA is given by:

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[\rho] = \int \mathrm{d}\mathbf{r}\ \epsilon_{\mathrm{xc}}^{\mathrm{GGA}}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})]\rho(\mathbf{r}), \tag{3.35}$$

where the GGA exchange-correlation energy density is constructed by modifying the $\epsilon_{\mathrm{xc}}^{\mathrm{LDA}}$ by an enhancement function F_{xc} [45]:

$$\epsilon_{\mathrm{xc}}^{\mathrm{GGA}}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})] = \epsilon_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})]F_{\mathrm{xc}}^{\mathrm{GGA}}[\rho(r), \nabla\rho(\mathbf{r})]. \tag{3.36}$$

There are many GGAs that have been developed over the years with different forms of the enhancement factor [10, 49–52]. The most well-known of these include the Perdew-Burke-Ernzerhof (PBE) functional [10] (the most highly cited paper in physics) which is used extensively within the physics community, and the Becke-Lee-Yang-Parr (BLYP) functional [52] used more heavily in computational chemistry. Despite belonging to the same category of exchange-correlation approximations, the PBE and BLYP functionals differ significantly in the rationale behind

their construction. The PBE functional is an example of a GGA that is is parameterised solely in terms of fundamental physical constants. The BLYP functional on the other hand is an empirically-fitted GGA, whose form was determined by fitting to results obtained from specific systems. These functionals generally offer an improvement over LDAs, however, the inclusion of additional terms that depend on the gradient electronic density does incur an increased computational cost.

3.5.3 Meta-GGA functionals

More complex functional approximations can be constructed by including other terms that depend on semi-local quantities in addition to the first derivative terms included in GGAs. These terms can include high-order derivatives of the density, or terms that depend on the kinetic energy density τ [53, 54]:

$$E_{\mathrm{xc}}^{\mathrm{mGGA}}[\rho] = \int \mathrm{d}\mathbf{r}\ G[\rho(\mathbf{r}), \nabla\rho(\mathbf{r}), \nabla^2\rho(\mathbf{r}), \tau(\mathbf{r}), \dots]. \tag{3.37}$$

By depending on more quantities, these meta-GGA functionals (mGGAs) can be constructed to satisfy even more properties than GGAs, however, they have not been shown to generalise consistently. In recent years, more advanced mGGAs have begun appearing, such as the SCAN functional [55], which has been shown to offer an improvement over GGAs and comparable performance to many hybrid functionals.

3.5.4 Hybrid functionals

Up to now, all of the classes of exchange-correlation functional approximations reviewed have resided firmly within the realms of the KS scheme of DFT. An extension to this scheme can be introduced by constructing new functionals from a mixture of standard KS functionals such as LDAs or GGAs and Hartree-Fock exchange. These functionals are known as hybrid functionals, whose general form is given by:

$$E_{\mathrm{xc}}^{\mathrm{HYB}} = (1-\alpha)E_{\mathrm{x}}^{\mathrm{GGA}}[\rho(\mathbf{r})] + \alpha E_{\mathrm{x}}^{\mathrm{HF}}[\{\phi_i(\mathbf{r})\}] + E_{\mathrm{c}}^{\mathrm{GGA}}[\rho(\mathbf{r})]. \tag{3.38}$$

Becke introduced the first of these hybrid functionals with a 50:50 mixture of LDA and HF, known as the "Becke half-half" functional [56]:

$$E_{\mathrm{xc}}^{\mathrm{HH}} = \frac{1}{2}E_{\mathrm{x}}^{\mathrm{HF}}[\{\phi_i(\mathbf{r})\}] + \frac{1}{2}E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho(\mathbf{r})]. \tag{3.39}$$

Becke justified this hybrid form by considering the limits of adiabatic connection integral:

$$E_{\mathrm{xc}} = \int_0^1 \mathrm{d}\lambda\ U_{\mathrm{xc},\lambda} \approx \frac{1}{2}U_{\mathrm{xc},0} + \frac{1}{2}U_{\mathrm{xc},1}. \tag{3.40}$$

Taking the average value of λ between these two limits provides a connection between the non-interacting KS system and the interacting system. At $\lambda = 0$, the system can be described by exchange-only Hartree-Fock theory, whereas at the limit of $\lambda = 1$, the system is fully interacting and correlation effects beyond that of Hartree-Fock must be included. To approximate the interacting limit, Becke chose to use the LDA functional, however, any GGA could equally have been used in its place with the same justification.

Hybrids have been shown to provide significant improvements over GGAs, for properties such as ionisation energies, electron affinities, and system geometries. Some of the most commonly used examples include the PBE0 functional, constructed from a 3:1 mixture of regular PBE and Hartree-Fock exchange [57], and the B3LYP functional, which uses a more advanced weighting of several LDA and GGA approximations alongside a HF contribution [9, 58]. One criticism of hybrid functionals is that due to the empirical fitting required in their construction, many hybrids may suffer from system bias, and like many mGGAs, may not generalise well to different types of systems. In fact, so many different hybrid functionals have been developed over the years that we now have a whole 'zoo' of functional approximations [8].

Since the development of the early hybrid functionals however, several additional arguments have been put forward which have provided non-empirical evidence in support of their construction, including an exact condition on the ratio of Hartree-Fock exchange to be used in combination with regular PBE [59]. Perdew and his co-workers also recently highlighted how the inclusion of a non-local exchange term

can dramatically improve predicted band gaps in solids [60]. Obviously, due to the inclusion of non-local Hartree-Fock exchange, hybrid functionals are significantly more computationally demanding than (semi-)local approximations.

3.5.5 The Generalised Kohn-Sham Scheme

In the previous subsection, it was demonstrated how Becke's adiabatic connection approach could be used to justify the existence of hybrid functionals. However, this argument is only valid when the exchange-correlation potential is constructed via an alternative approach to solving the KS equations known as the optimised effective potential (OEP) method, required for orbital-dependent functionals, which will be covered shortly. A rigorous justification for these functionals can instead be provided by the Generalised Kohn-Sham (GKS) scheme, of which the KS scheme is actually a specific case [61].

The main idea behind the KS scheme of DFT, is to find the optimal Slater determinant that represents a non-interacting N-electron system that yields the exact ground state electronic density as the fully interacting system. When attempting to calculate the expectation value of the Hamiltonian using this KS determinant, however, it is not possible to determine the total energy of the system, as the determinant cannot directly access the exchange-correlation energy. The main idea of GKS theory is to instead introduce an auxiliary system, still described by a single determinant, that takes into account some of the effects of the exchange-correlation contribution to the total energy.

Mathematically, the GKS scheme can be derived by introducing a generalised version of the universal functional, denoted by F_S:

$$F^S[\rho] = \min_{\Phi \to \rho} S[\Phi] = \min_{\{\phi_i\} \to \rho} S[\{\phi_i\}], \tag{3.41}$$

where $S[\Phi]$ is an energy expression dependent on $\Phi[\rho]$ or $\{\phi_i\}$ that obeys the following conditions:

1. The minimum of the functional $F^S[\rho]$ must exist and its functional derivatives with respect to ρ must be defined.

2. If we define the energy:

$$E^S[\{\phi_i\}, v_{\text{eff}}] = S[\{\phi_i\}] + \int \mathrm{d}\mathbf{r}\ v_{\text{eff}}(\mathbf{r})\rho(\mathbf{r}), \tag{3.42}$$

where $v_{\text{eff}}(\mathbf{r})$ is a multiplicative local potential, its minimum must yield a set of single-particle equations of the form:

$$\hat{O}^S[\{\phi_i\}]\phi_j + \hat{v}_{\text{eff}}\phi_j = \epsilon_i\phi_j, \tag{3.43}$$

where the operator $\hat{O}^S$ cannot depend explicitly on the potential $\hat{v}_{\text{eff}}$ and is invariant under unitary transformations to the set of single-particle orbitals.

To proceed, the total energy of the interacting system $E_0[v_{\text{en}}]$ is split into the total energy of the new GKS auxiliary system and a remainder term. By defining the difference between the universal functional $F[\rho]$ and the generalised functional $F^S[\rho]$ as a new functional $R^S[\rho] = F[\rho] - F^S[\rho]$, we can express $E_0[v_{\text{en}}]$ as:

$$\begin{aligned}
E_0[v_{\text{en}}] &= \min_{\rho\to N}\left\{F^S[\rho] + R^S[\rho] + \int \mathrm{d}\mathbf{r}\ v_{\text{en}}(\mathbf{r})\rho(\mathbf{r})\right\} \\
&= \min_{\rho\to N}\left\{\min_{\Phi\to\rho} S[\Phi] + R^S[\rho] + \int \mathrm{d}\mathbf{r}\ v_{\text{en}}(\mathbf{r})\rho(\mathbf{r})\right\} \\
&= \min_{\Phi\to N}\left\{S[\Phi] + R^S[\rho[\Phi]] + \int \mathrm{d}\mathbf{r}\ v_{\text{en}}(\mathbf{r})\rho[\Phi](\mathbf{r})\right\} \\
&= \min_{\{\phi_i\}\to N}\left\{S[\{\phi_i\}] + R^S[\rho[\{\phi_i\}]] + \int \mathrm{d}\mathbf{r}\ v_{\text{en}}(\mathbf{r})\rho[\{\phi_i\}](\mathbf{r})\right\}. \qquad (3.44)
\end{aligned}$$

Due to the second condition on the functional $S[\Phi]$, the single-particle GKS equations obtained from the minimisation 3.44 are:

$$\hat{O}^S[\{\phi_i\}]\phi_j + (\hat{v}_{\text{en}} + \hat{v}_R)\,\phi_j = \epsilon_j\phi_j, \tag{3.45}$$

where the potential $v_R(\mathbf{r})$ is defined by the functional derivative:

$$v_R(\mathbf{r}) = \frac{\delta R^S[\rho]}{\delta\rho(\mathbf{r})}. \tag{3.46}$$

The GKS equations can then be solved self-consistently like regular KS DFT equations.

We will now explore a couple of common choices of the functional $S[\Phi]$. Firstly, when $S[\Phi]$ is chosen to be equal to the kinetic energy functional for the determinant, i.e. $S[\Phi] = \langle\Phi|\,\hat{T}\,|\Phi\rangle$; the remainder functional will take the form:

$$R^S[\rho] = E_{\mathrm{H}}[\rho] + E_{\mathrm{x}}[\rho] + E_{\mathrm{c}}[\rho]. \tag{3.47}$$

Performing the minimisation 3.44 with this form of $R^S[\rho]$ then yields GKS equations that are the same as the standard single-particle equations of KS DFT:

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{en}}(\mathbf{r}) + v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{xc}}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \tag{3.48}$$

If we instead make the alternative choice of:

$$S[\Phi] = \langle\Phi|\,\hat{T} + \alpha\hat{V}_{\mathrm{ee}}\,|\Phi\rangle\,, \tag{3.49}$$

which gives:

$$R^S[\rho] = (1-\alpha)\,(E_{\mathrm{H}}[\rho] + E_{\mathrm{x}}[\rho]) + E_{\mathrm{c}}[\rho], \tag{3.50}$$

the minimisation yields the following single-particle equations:

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{en}}(\mathbf{r}) + v_{\mathrm{H}}(\mathbf{r}) + \alpha\hat{v}_{\mathrm{x}}^{\mathrm{HF}} + (1-\alpha)v_{\mathrm{x}}(\mathbf{r}) + v_{\mathrm{c}}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \tag{3.51}$$

The above GKS equations clearly define a functional that combines contributions from non-local Hartree-Fock exchange and some (semi-)local approximation; exactly the definition of a hybrid functional. The GKS scheme, therefore, provides a strong theoretical basis for the formation of this class of hybrid functionals. One can also derive the GKS equations for mGGAs, which are still non-local but do not include a term like Hartree-Fock exchange [60].

In summary, the GKS scheme essentially boils down to minimising over something other than the density, e.g. the KS orbitals. Depending on the functional, the equations that result from this alternative minimisation may or may not be the same as those obtained from minimising with respect to the density. LDA and PBE for example, yield the same equations, but the exact exchange functional (discussed in the next section) yields a different set of equations when the minimisation is performed with respect to the KS orbitals rather than the density.

3.5.6 Exact exchange

The last type of functional approximations to introduce are the exact-exchange (EXX) functionals, which look at first glance to be identical to those that include an amount of Hartree-Fock exchange:

$$E_{\mathrm{EXX}}[\{\phi_i[\rho]\}] = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\,\frac{\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})\phi_i(\mathbf{r}')\phi_j^*(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.52}$$

The subtle difference between the two is that for exact exchange functionals the orbitals experience a local potential. In the case of the GKS hybrid functionals, the single-particle equations that must be solved contain a non-local potential through the Hartree-Fock exchange potential operator.

As the EXX functional is an implicit/orbital-dependent functional, to obtain the exchange-correlation potential we must therefore use the OEP method, which was mentioned when discussing hybrid functionals. The details of the OEP method will be covered in § 3.6.

To obtain accurate results, EXX needs to be paired with an accurate correlation energy functional approximation. Unfortunately, using a fraction of EXX in combination with (semi-)local approximations for correlation yields worse results than one might expect. This is due to the cancellation of errors between the exchange and correlation parts of a typical (semi-)local functional, which does not occur when paired with EXX. In practice, EXX is usually used in combination with a correlation energy obtained via one of the correlated methods/post-HF mentioned in § 2.2.3, which introduces a dependency on the unoccupied KS orbitals and significantly drives up the cost of calculations. Due to the historical difficulties involved with the implementation of the OEP equation and their high computational demand, EXX-based functionals have not been used that often for practical applications. Thanks to current advancements, however, the development of these functionals has gained a lot of ground in recent years.

3.5.7 Spin-polarised DFT

Before concluding the discussion of exchange-correlation functionals, it is important to demonstrate how all of these classes of functionals can be extended to spin-polarised systems. At this stage, only spin-unpolarised systems have been considered, where each of the KS orbitals can be thought of as being doubly occupied (in a similar manner to restricted Hartree-Fock). The extension to systems with unpaired electrons is known as spin-DFT.

In spin-DFT, the behaviour of the spin-up and spin-down electrons is different across the two spin channels. Unlike in regular spin-unpolarised or restricted KS (RKS) theory, the system is governed not just by the total electronic density. In spin-DFT, it is instead controlled by the two spin densities, given by (for collinear spin cases):

$$\rho^{\sigma}(\mathbf{r}) = \sum_{i=1}^{N^{\sigma}} |\phi_i^{\sigma}(\mathbf{r})|^2, \ \sigma \in \{\uparrow, \downarrow\}. \tag{3.53}$$

An important question to answer is whether or not there exists an equivalent unique mapping between a spin-density and a spin-polarised potential. This query was addressed, following the same general ideas behind the original HK theorems, by von Barth and Hedin [62], who demonstrated the non-existence of such a mapping. In addition, Capelle and Vignale [63] and Eschrig and Pickett [64] showed that it is possible for several distinct potentials to have the same ground state density. Further investigations however have provided the proof of the uniqueness of the mapping [65, 66]; except for systems that are fully saturated to one particular spin-channel.

As the spin-densities are the fundamental variables of spin-DFT (not the total density), the minimisation of the kinetic energy functional must be performed differently than in RKS, by first searching for the determinant that yields given fixed spin-densities:

$$T_{\mathrm{s}}[\rho^{\uparrow}, \rho^{\downarrow}] = \min_{\Phi \to \rho^{\uparrow}, \rho^{\downarrow}} \langle \Phi | \hat{T} | \Phi \rangle . \tag{3.54}$$

Performing the minimisation yields the spin-polarised/unrestricted KS (UKS) equations:

$$\left[-\frac{\nabla^2}{2} + v^{\sigma}_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\mathbf{r}) + v^{\sigma}_{\text{xc}}(\mathbf{r})\right] \phi^{\sigma}_i(\mathbf{r}) = \epsilon^{\sigma}_i \phi^{\sigma}_i(\mathbf{r}), \tag{3.55}$$

where v^{σ}_{ext} is a spin-dependent potential representing both v_{en} and an external applied magnetic field, B_{ext}.

Notice that the UKS equations rely on the knowledge of the two local spin-potentials $v^{\uparrow}_{\text{xc}}(\mathbf{r})$ and $v^{\downarrow}_{\text{xc}}(\mathbf{r})$, calculated by the functional derivative of the exchange-correlation energy functional with respect to the spin-density. This is a necessary consequence of the unrestricted theory; that we do not have one single KS potential to represent the entire system.

In principle, if there is no external magnetic field acting on the system, the spin-DFT formalism should not be required, and standard DFT should give the same result. Nevertheless, in practice, open-shell systems (those with an unpaired electron) are usually treated with spin-DFT, not with DFT. One of the major developments presented in this thesis is a new formalism to treat these open-shell systems accurately whilst not invoking spin-DFT and to provide a previously unknown link between DFT and spin-DFT in this zero magnetic field limit; this will be introduced in § 5.

3.5.8 Ascending Jacob's Ladder

We now have explored the various classes of density functional approximations, discussing their development, relative accuracy, and computational complexity. As the field of DFT has matured, one of the main focuses has been making ever more accurate approximations for the exchange-correlation energy functional. A good way to visualise this development journey is through the 'Jacob's Ladder' of DFT [67], depicted in Figure (3.3). Functional approximations near the top of the 'ladder' are more complex but tend to provide improved accuracy. When ascending the 'rungs', the approximations introduce dependencies on more complex

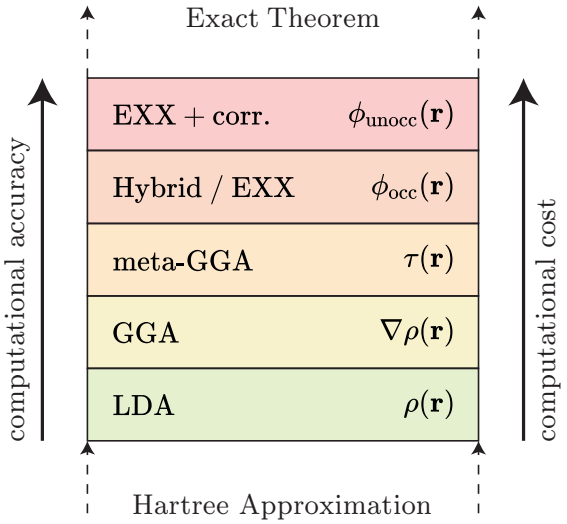


Figure 3.3: The 'Jacob's Ladder' of exchange-correlation functional approximations. The rungs of the ladder are coloured based on computational difficulty.

quantities, culminating in the dependency on the unoccupied orbitals in EXX-based methods paired with correlation obtained via post-HF methods.

3.6 Implicit functionals and the OEP method

We have mentioned several times now that in order to find the local exchange-correlation potential for implicit density functional approximations, we must make use of the optimised effective potential method (OEP).

There are multiple methods to derive the OEP equation, beginning with the original approach of Sharp and Horton who directly minimised the total energy of the Hartree-Fock energy [68] to obtain the equation. Although, the most common approach seen in the literature invokes the functional chain rule to directly construct the functional derivative of the total energy functional with respect to the density [69].

The OEP method can be derived by simply considering the one-to-one mapping between non-interacting density and the KS potential. By making a variation to

the potential $v_{\rm s}(\mathbf{r}) \rightarrow v_{\rm s}(\mathbf{r}) + \varepsilon\delta v_{\rm s}(\mathbf{r})$, the density must also change due to the one-to-one mapping between the two:

$$\delta\rho(\mathbf{r}) = \varepsilon \int \mathrm{d}\mathbf{r}' \; \chi[v_{\rm s}](\mathbf{r},\mathbf{r}')\delta v_{\rm s}(\mathbf{r}'), \tag{3.56}$$

where the object $\chi[v_{\rm s}](\mathbf{r},\mathbf{r}')$ is known as the density-density response function:

$$\chi[v_{\rm s}](\mathbf{r},\mathbf{r}') = \frac{\delta\rho[v_{\rm s}](\mathbf{r})}{\delta v_{\rm s}(\mathbf{r}')} = \frac{\delta}{\delta v_{\rm s}(\mathbf{r}')}\left\{\sum_{i=1}^{N}|\phi_i[v_{\rm s}](\mathbf{r})|^2\right\} = \sum_{i=1}^{N}\phi_i^*[v_{\rm s}](\mathbf{r})\frac{\delta\phi_i[v_{\rm s}](\mathbf{r})}{\delta v_{\rm s}(\mathbf{r}')} + \text{c.c.} \tag{3.57}$$

The functional derivative of the orbitals with respect to the potential can then easily be found using first-order perturbation theory:

$$\frac{\delta\phi_i[v_{\rm s}](\mathbf{r})}{\delta v_{\rm s}(\mathbf{r}')} = \sum_{k\neq i}\frac{\phi_i[v_{\rm s}](\mathbf{r}')\phi_k^*[v_{\rm s}](\mathbf{r}')\phi_k[v_{\rm s}](\mathbf{r})}{\epsilon_i[v_{\rm s}] - \epsilon_k[v_{\rm s}]}. \tag{3.58}$$

Separating the sum over $i \neq k$ into two sums over occupied, i, and unoccupied states, a, respectively, then allows the response function to be written as:

$$\chi[v_{\rm s}](\mathbf{r},\mathbf{r}') = -\sum_{i=1}^{N}\sum_{a=N+1}^{\infty}\frac{\phi_i^*[v_{\rm s}](\mathbf{r})\phi_a[v_{\rm s}](\mathbf{r})\phi_i[v_{\rm s}](\mathbf{r}')\phi_a^*[v_{\rm s}](\mathbf{r}')}{\epsilon_a[v_{\rm s}] - \epsilon_i[v_{\rm s}]} + \text{c.c.} \tag{3.59}$$

Notice that we are now writing quantities as explicit functionals of the potential, not the density. Also, due to the denominator, the response function is defined to be negative definite. We can also write other quantities such as the exchange-correlation functional in this manner. However, to express these quantities as either functionals of the density or the potential we must require that for $v_{\rm s} = v_{\rm s}[\rho]$ the values of the explicit and implicit functionals of the density must be equal, i.e. $\phi_i[v_{\rm s}[\rho]] = \phi_i[\rho]$, $E_{\rm xc}[v_{\rm s}[\rho]] = E_{\rm xc}[\rho]$, etc.

Turning our attention back to the exchange-correlation functional, for a change in the density $\rho(\mathbf{r}) \rightarrow \rho(\mathbf{r}) + \varepsilon\delta\rho(\mathbf{r})$ induced by the change in the potential, the xc-energy will change by (for small ε):

$$\Delta E_{\rm xc} = \varepsilon\int \mathrm{d}\mathbf{r} \; \delta\rho(\mathbf{r})\frac{\delta E_{\rm xc}[\rho]}{\delta\rho(\mathbf{r})} = \varepsilon\int \mathrm{d}\mathbf{r}' \; \delta v_{\rm s}(\mathbf{r}')\int \mathrm{d}\mathbf{r}\chi[v_{\rm s}[\rho]](\mathbf{r},\mathbf{r}')\frac{\delta E_{\rm xc}[\rho]}{\delta\rho(\mathbf{r})}. \tag{3.60}$$

By considering how the xc-functional changes directly as a consequence of variations in the potential we can also write that:

$$\Delta E_{\rm xc} = \varepsilon\int \mathrm{d}\mathbf{r}' \; \delta v_{\rm s}(\mathbf{r}')\frac{\delta E_{\rm xc}[v_{\rm s}]}{\delta v_{\rm s}(\mathbf{r}')}\bigg|_{v_{\rm s}=v_{\rm s}[\rho]}. \tag{3.61}$$

Equating these two expressions yields:

$$\varepsilon \int \mathrm{d}\mathbf{r}'\ \delta v_{\mathrm{s}}(\mathbf{r}') \int \mathrm{d}\mathbf{r} \chi[v_{\mathrm{s}}[\rho]](\mathbf{r},\mathbf{r}') \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(\mathbf{r})} = \varepsilon \int \mathrm{d}\mathbf{r}'\ \delta v_{\mathrm{s}}(\mathbf{r}') \frac{\delta E_{\mathrm{xc}}[v_{\mathrm{s}}]}{\delta v_{\mathrm{s}}(\mathbf{r}')}\bigg|_{v_{\mathrm{s}}=v_{\mathrm{s}}[\rho]}, \quad (3.62)$$

which by inspection implies:

$$\int \mathrm{d}\mathbf{r} \chi[v_{\mathrm{s}}[\rho]](\mathbf{r},\mathbf{r}') \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(\mathbf{r})} = \frac{\delta E_{\mathrm{xc}}[v_{\mathrm{s}}]}{\delta v_{\mathrm{s}}(\mathbf{r}')}\bigg|_{v_{\mathrm{s}}=v_{\mathrm{s}}[\rho]}. \quad (3.63)$$

Using the standard definition of the exchange-correlation potential from Equation 3.26, we can write the OEP equation as:

$$\int \mathrm{d}\mathbf{r}\ \chi[v_{\mathrm{s}}[\rho]](\mathbf{r},\mathbf{r}') v_{\mathrm{xc}}[\rho](\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[v_{\mathrm{s}}]}{\delta v_{\mathrm{s}}(\mathbf{r}')}\bigg|_{v_{\mathrm{s}}=v_{\mathrm{s}}[\rho]}. \quad (3.64)$$

The xc-potential can then be determined by inverting the OEP equation:

$$v_{\mathrm{xc}}[\rho](\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \chi^{-1}[v_{\mathrm{s}}[\rho]](\mathbf{r},\mathbf{r}') \frac{\delta E_{\mathrm{xc}}[v_{\mathrm{s}}]}{\delta v_{\mathrm{s}}(\mathbf{r}')}\bigg|_{v_{\mathrm{s}}=v_{\mathrm{s}}[\rho]}. \quad (3.65)$$

One important point to highlight is that the potential obtained from the OEP equation is only determined up to a constant. So for a potential $v_{\mathrm{xc}}[\rho](\mathbf{r})$ that is a solution to Equation 3.64, any potential $v_{\mathrm{xc}}[\rho](\mathbf{r}) + c$ is also a valid solution. This draws a nice parallel to the HK theorems where the potential is also shown to be determined only up to a constant. More explicitly, a constant falls within the null space of the response function, such that:

$$\int \mathrm{d}\mathbf{r}\ \chi[v_{\mathrm{s}}[\rho]](\mathbf{r},\mathbf{r}') c = 0. \quad (3.66)$$

The OEP equation does not only apply to the exchange-correlation potential. In fact, the derivation above can be applied to any density functional. The most commonly used case of the OEP equation is for the exact-exchange functional. Using the definition of the EXX energy results in:

$$\frac{\delta E_{\mathrm{EXX}}[v_{\mathrm{s}}]}{\delta v_{\mathrm{s}}(\mathbf{r}')}\bigg|_{v_{\mathrm{s}}=v_{\mathrm{s}}[\rho]} = -\sum_{i=1}^{N} \sum_{a=N+1}^{\infty} \iint \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{x} \frac{\rho(\mathbf{x},\mathbf{r})}{|\mathbf{r}'-\mathbf{x}|} \frac{\phi_i^*(\mathbf{r}')\phi_i(\mathbf{x})\phi_a(\mathbf{r}')\phi_a^*(\mathbf{r})}{\epsilon_a - \epsilon_i} + \text{c.c.}, \quad (3.67)$$

where for compactness the explicit functional dependencies of the orbitals on the potential and the density have been dropped. Using this form for the right-hand

side of the OEP equation we then have knowledge of all of the quantities required to find its solution.

In its current form above, the OEP equation is well-known to be costly and difficult to solve numerically. The perceived high-cost of the OEP equation is mainly due to the unnecessarily expensive methods that have traditionally been used to solve it; where the auxiliary basis is fixed in size and the orbital basis is increased until convergence is reached [70, 71]. In order to reach convergence, the space spanned by the orbital basis must completely cover the space of the auxiliary basis, which, in practice, requires huge orbital basis sets even for simple systems. The OEP equation is also subject to numerical instabilities that can result in some undesirable potentials, which in the past has made performing reliable calculations difficult.

Two commonly used approximations that can be used to completely bypass this problems, are the Krieger, Li, and Iarfrate (KLI) approximation [72, 73], and the more accurate (but related) common energy denominator approximation (CEDA) [74, 75].

In CEDA, the energy difference $\epsilon_a - \epsilon_i$ between every pair of occupied-unoccupied orbitals is approximated by the average value of the differences, Δ_{CEDA}, by invoking the Unsöld approximation [76]. As all of these differences are positive, Δ_{CEDA} is strictly a positive quantity. This enables the left-hand side of the OEP equation to be written as:

$$\int \mathrm{d}\mathbf{r}\ \chi(\mathbf{r},\mathbf{r}')v_{\text{xc}}(\mathbf{r}') = -\frac{2}{\Delta_{\text{CEDA}}}\Bigg[\int \mathrm{d}\mathbf{r}\ \rho(\mathbf{r}')v_{\text{xc}}(\mathbf{r}') - \sum_{i=1}^{N}\sum_{j=1}^{N}\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})\int \mathrm{d}\mathbf{r}'\phi_i^*(\mathbf{r}')\phi_j^*(\mathbf{r}')v_{\text{xc}}(\mathbf{r}')\Bigg], \quad (3.68)$$

where the completeness relation has been used to remove the sum over the unoccupied orbitals:

$$\sum_{a=N+1}^{\infty}\phi_a(\mathbf{r})\phi_a^*(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') - \sum_{j=1}^{N}\phi_j(\mathbf{r})\phi_j^*(\mathbf{r}'). \quad (3.69)$$

In the KLI approximation, the double summation is instead performed over occupied orbitals and then by all of the orbitals (not just the unoccupied ones). This,

therefore, means that the sign associated with each of the individual energy differences is not fixed, leading to a less accurate representation of the response function. The response function in the KLI approximation is therefore not a strictly negative operator as it should be, unlike in CEDA.

The CEDA approximation completely bypasses the numerical instabilities that affect the full solution to the OEP equation and can be used as a robust alternative to the full solution. The next chapter will explore these issues further and discuss how we can alleviate them.

3.7 Implementation for molecular systems

Having reviewed all of the analytical theory required for understanding ground state DFT, what remains unanswered is how to put these ideas into practice. When it comes to implementing the KS equations, there are two main approaches:

1. Represent the KS orbitals directly on a real-space grid.
2. Represent the KS orbitals using a basis set.

Performing calculations on a real-space grid is the more simple approach, with the most important concern revolving around how fine a grid to use. DFT calculations often require a rather fine grid, which has the knock-on effect of increasing the computational cost of calculations. As a result, grid-based methods are generally only used for small systems. One should not completely disregard this approach, however, as modern supercomputers allow for calculations to be parallelised across huge numbers of cores.

The far less demanding approach is to use a basis set to represent the KS orbitals. For molecular/finite systems, the most commonly used basis sets are linear combinations of atomic orbitals (LCAOs), which are adopted for all calculations presented in this thesis. Within this approach, the KS orbitals are written as a

weighted sum over all of the basis functions in the orbital basis set:

$$\phi_i(\mathbf{r}) = \sum_k^{n_{\text{orb}}} c_{ik}\xi_k(\mathbf{r}). \tag{3.70}$$

These localised basis sets work well for molecular systems as the electronic orbitals are generally localised near the atoms of the molecule. For periodic systems, plane waves are far more commonly used instead, due to the presence of often strongly delocalised electrons in, for example, metallic systems.

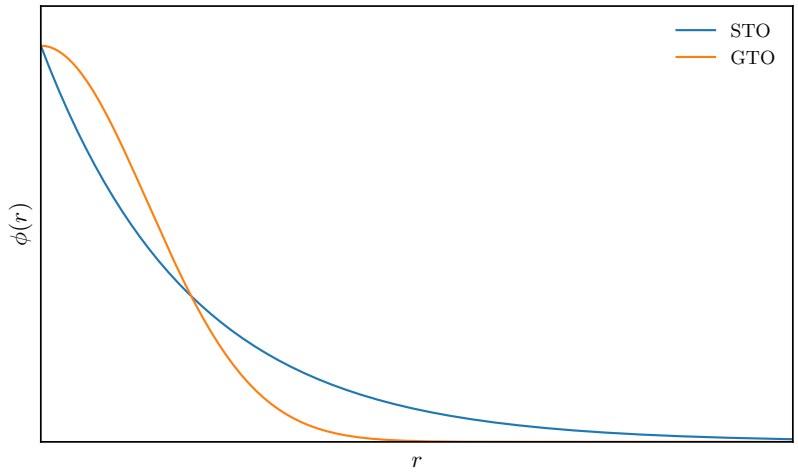


Figure 3.4: The shape of the radial part of a STO and a GTO with identical parameters. The STO shows an exponential decay at large r and a cusp at $r \to 0$.

A logical choice of LCAOs are Slater-type orbitals (STOs) [77] of the form (in Cartesian coordinates):

$$\xi^{\text{STO}}_{k,a,b,c}(\mathbf{r}) = N_k x^a y^b z^c e^{-\zeta_k r}, \tag{3.71}$$

where a, b, and c are integers that control the angular momentum $l = a + b + c$, N_k is a normalisation constant, and ζ_k is a constant related to the screened nuclear charge that controls the width of the orbital. STOs are similar to the hydrogen-like solutions of the Schrödinger equation; however, they do not possess any radial nodes and are not true spherical harmonics.

Slater orbitals are a particularly suitable choice of functions to represent the atomic orbitals, as they correctly decay exponentially at large r and satisfy Kato's cusp

condition [78, 79] at the point of the nucleus. However, the use of STOs is usually limited in calculations involving atomic species or linear molecules because of the difficulty in calculating the multi-electron integrals required for any multi-atomic system.

To bypass this difficulty, Gaussian-type orbitals (GTOs) [80–82] can be used instead, which in Cartesian coordinates take the form:

$$\xi^{\mathrm{GTO}}_{k,a,b,c}(\mathbf{r}) = N_k x^a y^b z^c e^{-\alpha_k r^2}, \tag{3.72}$$

where again the angular momentum of the orbital is given by $a + b + c = l$, N_k is a normalisation constant, and the constant α_k plays a similar role to ζ_k in STOs. The difference in the general shape of a STO and a GTO is shown in Figure 3.4.

As the product of two GTOs centered at different atoms is simply another GTO centered between the two atoms, the multi-electron integrals can be calculated analytically. This is a major computational advantage of GTOs over STOs. However, due to the functional form of GTOs they do not exhibit the correct behaviour near to the centres of the nuclei. To combat this, fixed linear combinations of individual or 'primitive' GTOs are used, known as 'contracted' GTOs (CGTOs); given by:

$$\xi^{\mathrm{CGTO}}_{k,a,b,c}(\mathbf{r}) = \sum_{l}^{n_{\mathrm{k}}} c_l \xi^{\mathrm{GTO}}_{l,a,b,c}(\mathbf{r}) = N_k \sum_{l}^{n_{\mathrm{k}}} c_l x^a y^b z^c e^{-\alpha_l r^2}, \tag{3.73}$$

where the coefficients c_l are called the contraction coefficients. By combining GTOs with a range of values for α_l, and weighting them by the contraction coefficients the resulting CGTOs are able to achieve a small r behaviour closer to that of a STO. Although, even with an infinite number of GTOs included in the contraction, the resulting CGTO will always have a zero gradient at the nuclear centre.

The simplest class of GTO basis sets are minimal basis sets, where only one CTGO is used to represent each atomic orbital in the system. The most well-known of these are the STO-nG basis sets developed by John Pople [83]. Despite being extremely cheap computationally, minimal basis sets are not very accurate, and thus larger basis sets with multiple CGTOs per orbital are used in modern calculations.

To construct these larger basis sets, each atomic orbital is allowed to be represented by multiple CTGOs. By including two CGTOs per orbital a double-zeta basis set is constructed, with three per orbital a triple-zeta basis set, and with four a quadruple-zeta basis set. This number can be extended further to construct even larger basis sets, known as n-zeta basis sets where n is the number of CGTOs per atomic orbital. One can also construct split-valence basis sets, where only one CGTO is used to represent core atomic orbitals, and then multiple CGTOs for the valence orbitals. This type of basis set was also pioneered by Pople and his co-workers [84].

On top of the fitted CGTOs, the majority of modern basis sets include polarisation functions. These functions provide a better representation of systems with orbitals that polarise to one side of their respective atoms as they are brought into close proximity with those from other atoms. In the presence of a p-orbital for example, an s-orbital will often polarise in the direction of the p-orbital. To include polarised functions, CGTOs with a higher angular momentum than those that are occupied are incorporated into the basis set. Due to the inclusion of polarisation functions, these basis sets are known as polarised basis sets.

Another type of function that are often added are diffuse functions. These are CGTOs with very small values for the exponent coefficients, allowing for the representation of electrons that are bound far from the atomic nucleus. These functions are essentially a requirement for calculations involving anions or highly electronegative atoms, e.g. fluorine, and not including these functions can cause computed properties for these systems to vary dramatically. Basis sets with these diffuse functions added are called augmented basis sets.

The construction of basis sets is always a balance between accuracy and computational efficiency. In theory, as more basis functions are added to the orbital basis set, we should gradually converge to the exact representation of the real orbitals of the system; with full convergence reached in the infinite basis set limit. Some basis sets are specifically designed with this convergence in mind, the main example being

Dunning's correlation-consistent basis sets [85, 86]. These basis sets are compactly written as cc-pVnZ, which stands for correlation-consistent, polarised valence, n-zeta, with n again representing the number of CGTOs per atomic orbital. The augmented versions of these sets are simply written by adding the prefix aug- to the base acronym. As these basis sets are designed to converge systematically with basis set size, they are used frequently for DFT calculations and are the choice of basis sets for every calculation presented in this thesis.

Having decided on using Gaussian-type orbitals to represent the atomic orbitals, attention reverts to solving the KS equations. In fact, the SCF procedure is the same as for solving the HF equations (2.25), with the elements of the Fock matrix (now termed the KS Fock matrix) given by:

$$F_{kl} = F_{kl}^{\text{core}} + F_{kl}^{\text{H}} + F_{kl}^{\text{xc}}, \tag{3.74}$$

where the core and Hartree components are given by:

$$F_{kl}^{\text{core}} = \int \mathrm{d}\mathbf{r} \xi_k(\mathbf{r}) \left[-\frac{\nabla^2}{2} + v_{\text{en}}(\mathbf{r}) \right] \xi_l(\mathbf{r}), \tag{3.75}$$

$$F_{kl}^{\text{H}} = \sum_{m=1}^{n_{\text{orb}}} \sum_{n=1}^{n_{\text{orb}}} \rho_{mn} \iint \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{r}' \, \frac{\xi_k(\mathbf{r})\xi_l(\mathbf{r})\xi_m(\mathbf{r}')\xi_n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{3.76}$$

and ρ_{mn} is the density matrix $\sum_{i=1}^{N} c_{im} c_{in}$.

The calculation of F_{kl}^{xc} is not however quite as straightforward, owing to the many different types of exchange-correlation functionals that can be used within a DFT calculation. Normally, this part of the Fock matrix is constructed separately on a grid, transformed into its basis set representation, and then finally combined with the core and Hartree terms:

$$F_{kl}^{\text{xc}} = \sum_{g=1}^{n_{\text{grid}}} w_g \xi_k(r_g) v_{\text{xc}}(r_g) \xi_l(r_g), \tag{3.77}$$

where the coefficient w_g is the weighting of the grid at r_g. As most of the electronic density can be found within the vicinity of the atoms, the grid can be constructed with this in mind to increase the computational efficiency of this costly step in the SCF cycle [87].

3.7.1 Convergence methods

All of the electronic structure methods that we have discussed are based on solving an iterative procedure that aims to minimise the total energy of the system in question. In the case of KS theory, to begin any calculation one must choose an initial guess for the orbitals to enter the KS equations for the first SCF cycle, to generate a new guess of the orbitals. The total energy is then constructed using the new guess and compared to the previous total energy. If the new total energy falls within a certain tolerance of the previous total energy, the calculation is said to then be converged. However, performing successive iterations in this way often results in the build-up of oscillations between successive solutions, and the electronic density never reaches convergence. It is important to invoke additional convergence techniques in order to avoid the build-up of these oscillations.

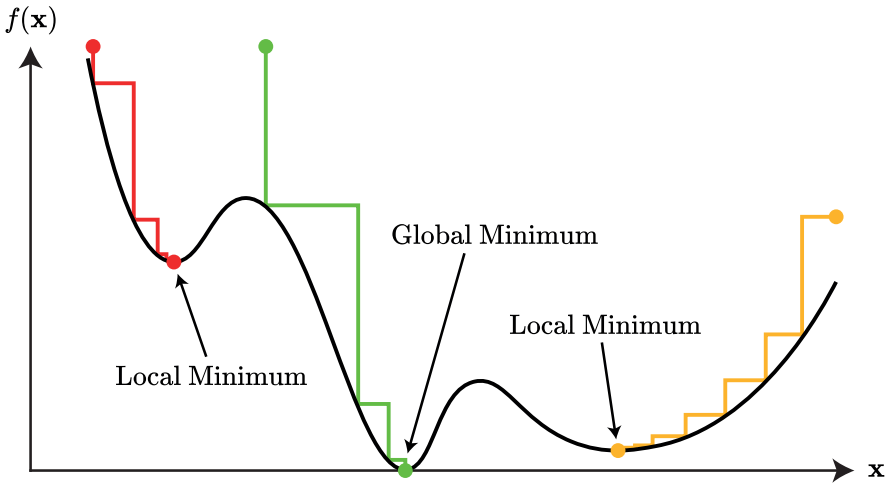


Figure 3.5: Comparison between local and global minima. For each of the different starting guesses, a different result is obtained after convergence due to the complexity of the functions landscape.

Even with the more complex techniques described later in this section, reaching a converged solution does not necessarily mean we have successfully found the global minimum in the space of all possible Slater determinants. It is common for there to

be several local minima present within the energy landscape. This is demonstrated in Figure 3.5, where certain minima (e.g. the red minimum) are clearly far from the true lowest value of the function, others (e.g. the orange minimum) are close but still not the true global minimum; both are converged solutions, however.

3.7.1.1 Density mixing

The simplest method that can be used to try to dampen density oscillations near convergence is density mixing, where the KS Fock matrix for the current iteration in the SCF procedure is given by a linear weighting between the current and previous matrices:

$$F_{kl}^{n} = \alpha F_{kl}^{n} + (1 - \alpha) F_{kl}^{n-1}. \tag{3.78}$$

For small values of α, convergence is slower but more robust, as less of the new Fock matrix is mixed with the previous one. With large values of α, convergence is faster but prone to harsh oscillations in the density and overshooting. In practice, a value of α around 0.2 is a good balance between SCF efficiency and stability for a wide variety of systems.

3.7.1.2 Direct inversion in the iterative subspace (DIIS)

In more stubborn cases, relying on a simple mixing of the Fock matrix is not sufficient. An alternative scheme was developed by Pulay, known as the direct inversion in the iterative subspace (DIIS) routine [88, 89]. The DIIS procedure makes the assumption that the final solution to an iterative problem $\mathbf{p}^{f}$ can be expressed approximately as a linear combination of the previous m iterative guesses:

$$\mathbf{p} = \sum_{i=1}^{m} c_i \mathbf{p}^{i}. \tag{3.79}$$

In practice, not all of the previous trial vectors are used in the expansion, such that the above summation runs over the previous n_{DIIS} iterations. The coefficients of

this expansion are determined by defining a set of error vectors $\Delta\mathbf{p}^i = \mathbf{p}^{i+1} - \mathbf{p}^i$, which collectively define the error associated with the vector $\mathbf{p}$:

$$\Delta\mathbf{p} = \sum_{i=1}^{m} c_i \Delta\mathbf{p}^i, \tag{3.80}$$

which are minimised under the constraint that the set of coefficients $\{c_i\}$ sum to one. The intuition behind this constraint can be seen by writing each of the trial solutions $\mathbf{p}^i$ as the sum of the exact solution $\mathbf{p}^f$ and an error vector $\mathbf{e}^i$, so that the DIIS solution can be written as:

$$\mathbf{p} = \mathbf{p}^f \sum_{i=1}^{m} c_i + \sum_{i=1}^{m} c_i \mathbf{e}^i. \tag{3.81}$$

To minimise the absolute error, the error term must be zero and $\mathbf{p} = \mathbf{p}^f$, therefore arriving at the condition that $\sum_{i=1}^{m} c_i = 1$. The minimisation is performed by constructing the following Lagrangian:

$$\mathcal{L} = \sum_{i,j=1}^{m} c_j B_{ji} c_i - 2\lambda \left(\sum_{i=1}^{m} c_i - 1 \right), \tag{3.82}$$

where $B_{ij} = \langle \Delta\mathbf{p}^i | \Delta\mathbf{p}^i \rangle$, and setting its partial derivative with respect to the coefficient c_k equal to zero:

$$\frac{\partial \mathcal{L}}{\partial c_k} = \sum_{j=1}^{m} c_j B_{kj} + \sum_{i=1}^{m} c_i B_{ik} - \lambda = 2 \sum_{i=1}^{m} c_i B_{ki} - \lambda = 0. \tag{3.83}$$

This equation can then be inverted to find the values of the set of coefficients for the DIIS solution. For the case of KS DFT, we can apply the DIIS routine in the construction of the Fock matrix and use the same method as outlined above.

3.7.1.3 Maximum overlap method (MOM)

When dealing with nearly or fully degenerate systems, i.e. where the energy difference between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) is very small or zero, the DIIS algorithm presented above is no longer able to converge; except for systems where the highest-occupied KS orbital is a singly occupied s-orbital ($l = 0$). Within the RKS scheme,

every open-shell system falls into this category of systems, as well as molecules that are stretched beyond their equilibrium geometries. As a calculation involving one of these systems approaches convergence, it will begin to oscillate between occupying each of the degenerate KS orbitals, and thus convergence is never reached.

An algorithm called the maximum overlap method (MOM) [90] can be used to deal with these cases. The idea of this method is that for each SCF iteration, the newly occupied orbitals should be those that have the largest overlap with the occupied orbitals from the previous iteration. The amount of overlap can be measured by calculating the projection p_i of the new i-th orbital onto the previous set of occupied orbitals:

$$p_i = \sum_{k=1}^{n_{\text{orb}}} \sum_{l=1}^{n_{\text{orb}}} \sum_{j=1}^{N} \rho_{jk}^{n} S_{kl} \rho_{li}^{n-1}. \tag{3.84}$$

The new occupied orbitals are then the N with the largest values in the set of $\{p_i\}$. The MOM method generally leads to converged solutions that do not obey the Aufbau principle, which at first may seem somewhat alarming. However, as the KS orbitals are not the physical electrons this solution can safely be accepted.

3.7.1.4 Gradient minimisation methods

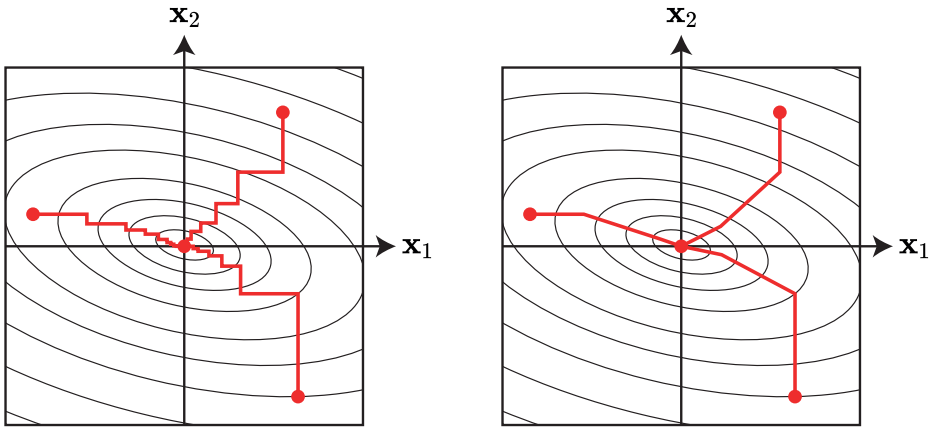


Figure 3.6: A schematic representation of the efficiency of steepest descent vs conjugate gradient for different starting guesses.

To find the minimum of any function or functional, we must be able to plot a route from our starting guess to its minimum efficiently and robustly. There are several schemes that can be used, but perhaps the most simple is known as the steepest descent method, which provided that the gradient of a function f can be defined, chooses the direction to take at each iterations starting point x_n to be the direction of steepest descent, i.e. $-f'(x_n)$. The new starting point for the next iteration is therefore given by:

$$x_{n+1} = x_n - \alpha_n f'(x_n), \tag{3.85}$$

where α_n is determined from the minimum of $f(x_{n+1})$. The steepest descent approach is prone to a couple of issues, especially when certain parameters of a function are highly correlated with each other. In these cases, the algorithm may take a lot of steps to reach the minimum, slowing down the computation. This can be seen in the left subplot of Figure 3.6, where it takes the steepest descent algorithm many steps when starting in the shallow regions of the functional landscape.

An alternative and improved method for plotting a route to the minimum can be obtained by slightly modifying the steepest descent algorithm after the first iteration, to instead step in a direction that is conjugate to the direction of the previous iteration. This method is known as the conjugate gradient approach, and it can provide far superior performance over the steepest descent method; especially in the cases where the latter tends to struggle (as shown in Figure 3.6).

3.7.2 Codes Used

All of the major developments of this thesis were implemented into the Gaussian basis set code HIPPO* [91]. The required one-electron and two-electron integrals for the Cartesian Gaussian basis functions are calculated using the quantum chemistry package GAMESS [92, 93]. This code utilises the LIBXC library for exchange-correlation functionals [94] to obtain the xc-potential and energy density

*Contact N. Lathiotakis at lathiot@eie.gr for information.

on a numerical grid. Basis set data was extracted from the Basis Set Exchange (BSE) database [95–97]. Additionally, the open-source code PSI4 [98, 99] was used to perform any unrestricted Hartree-Fock/DFT calculations, as well as any coupled cluster calculations that are not supported by the HIPPO code.

Chapter 4

Solving the OEP equation & constrained density functional approximations

This chapter will discuss how the OEP equation can be implemented and solved in a finite basis set code, and demonstrate that despite well-known mathematical and numerical issues that have plagued the OEP method for finite basis implementations in the past numerically robust solutions can in fact be obtained.

The OEP equation will be solved using the constrained minimisation method, first proposed by Gidopoulos and Lathiotakis [100, 101], which applies physically intuitive constraints to the effective potential to correct the asymptotic behaviour of the KS potential. I rewrote the existing routines that implemented this method in the HIPPO code to improve their computational efficiency and use-ability, and present results highlighting the systematic improvements this method provides for several system properties.

The major improvement of the constrained method presented in this chapter however is the study of when the positivity constraint can be relaxed. This constraint will be introduced and explained in a following section and previous implementations of the method enforced it on the OEP potential. This constraint proved to

be a computational bottleneck, and lifting it allows the constrained method to be applied efficiently to larger systems.

This chapter is based on work published in Ref. [102].

4.1 Theory

4.1.1 The asymptotic behaviour of the Kohn-Sham potential

A key analytical property of the exact KS Hxc potential for a finite N-electron system is that its asymptotic behaviour far from the system should decay $\propto \frac{N-1}{r}$. It is also known that the Hartree potential alone should decay as $\propto \frac{N}{r}$. Thus, is it easy to deduce the exact condition on the asymptotic behaviour of the exact KS xc potential, as $\propto \frac{-1}{r}$. This is not however the asymptotic behaviour found when considering standard density functional approximations (DFAs), which model the KS Hxc potential as having a decay $\propto \frac{N}{r}$. The cause of this incorrect behaviour is self-interaction. All of the most commonly used (semi-)local density functional approximations, e.g. LDA and GGAs, do not correct for self-interaction errors (and those that do only partially correct these errors).

It is possible to gauge the strength of the self-interaction errors inherent in a given density functional approximation by expressing the effective potential in terms of an effective screening density, ρ_{scr}. Mathematically this connection can be written as a Poisson equation [100, 103]:

$$\nabla^2 v_{\text{eff}}(\mathbf{r}) = -4\pi\rho_{\text{scr}}(\mathbf{r}), \tag{4.1}$$

which may be solved using Green's theorem; yielding:

$$v_{\text{eff}}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \frac{\rho_{\text{scr}}(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}. \tag{4.2}$$

This relation allows us to quantify the strength of self-interactions by integrating the screening density over all space to calculate the screening charge, Q_{scr}, and

checking its value against the exact $N-1$ charge expected. If $Q_{\text{scr}} = N-1$, then the functional can be thought of as self-interaction free or corrected (at least with regards to the asymptotic behaviour, as these interactions can manifest in other ways). If however, $Q_{\text{scr}} > N-1$, self-interactions are to some degree present in the underlying density functional approximation.

Obtaining an accurate representation of the decay of the exact potential far from the system is in fact a key concern when one is interested in, for example, ionisation energies. Koopmans' theorem, where the ionisation energy of the i-th electron is equal to minus the eigenvalue of i-th electronic orbital (under the frozen orbital approximation) was already discussed earlier in § 2.2.2.1. This relationship between eigenvalue and ionisation energy applies to every occupied electron within HF theory. However, within DFT no such general relation exists, except for the proof for the exact equality between minus the HOMO eigenvalue and the principal ionisation energy (commonly termed the Koopmans' theorem of DFT) [104, 105]. It has been shown recently however that there does exist an approximate correspondence between each occupied KS eigenvalue and their respective ionisation energies [106–108].

Analytically, the asymptotic decay of electronic density of the interacting system can be shown to have an explicit dependence on the exact ionisation energy [109]:

$$\lim_{\mathbf{r}\to\infty} \rho(\mathbf{r}) \propto \exp\Big\{-2\sqrt{2I_N}|\mathbf{r}|\Big\}. \tag{4.3}$$

A similar expression can also be derived for the non-interacting KS system, where the decay of the density depends on the HOMO eigenvalue [110]:

$$\lim_{\mathbf{r}\to\infty} \rho(\mathbf{r}) \propto \exp\Big\{-2\sqrt{-2\epsilon_N}|\mathbf{r}|\Big\}. \tag{4.4}$$

By construction however, the KS density is equal to the exact interacting density, and therefore the two above relations can be equated, resulting in $I_N = -\epsilon_N$. It is important to note that although this relation is exact in the case of the exact KS potential, it becomes approximate with the introduction of density functional

approximations. Consequently, it is possible to use the accuracy of calculated ionisation energies as an example to benchmark DFAs against one another (see Table 4.3).

4.1.2 The OEP equation for finite basis sets

Solving the KS equations for an orbital-dependent functional requires the use of the OEP equation, which cannot simply be constructed on the grid and fed into the construction of the Fock matrix. The following discussion focuses on how the OEP equation can be solved numerically when the orbitals and effective potential are expanded in (different) finite basis sets.

We begin with the general OEP equation:

$$\int \mathrm{d}\mathbf{r}'\chi(\mathbf{r},\mathbf{r}')v_{\mathrm{eff}}(\mathbf{r}') = b(\mathbf{r}), \tag{4.5}$$

which is not restricted to the specific case of $v_{\mathrm{eff}} = v_{\mathrm{xc}}$. Here, the effective potential and $b(\mathbf{r})$ depend on the quantity of interest; this is usually either the xc potential or the Hartree-xc (Hxc) potential.

There are several implementations of the OEP equation for finite systems, all of which begin by expanding the effective potential in terms of an effective screening density, $\rho_{\mathrm{scr}}(\mathbf{r})$, which is then, in turn, expanded in terms of an auxiliary basis set:

$$v_{\mathrm{eff}}(\mathbf{r}') = \int \mathrm{d}\mathbf{x}\frac{\rho_{\mathrm{scr}}(\mathbf{x})}{|\mathbf{x}-\mathbf{r}'|}; \quad \rho_{\mathrm{scr}}(\mathbf{r}) = \sum_{k}^{n_{\mathrm{aux}}} \rho_k^{\mathrm{s}}\theta_k(\mathbf{r}), \tag{4.6}$$

where $\{\theta_k\}$ is the set n_{aux} auxiliary basis functions. Inserting this form for the effective potential into the OEP equation yields:

$$\sum_{k}^{n_{\mathrm{aux}}} \rho_k^{\mathrm{s}} \iint \mathrm{d}\mathbf{r}'\,\mathrm{d}\mathbf{x}\frac{\chi(\mathbf{r},\mathbf{r}')}{|\mathbf{r}'-\mathbf{x}|}\theta_k(\mathbf{x}) = b(\mathbf{r}). \tag{4.7}$$

Defining the function $\tilde{\theta}_k(\mathbf{r}') = \int d\mathbf{x}\theta_k(\mathbf{x})/|\mathbf{r}'-\mathbf{x}|$, multiplying by $\tilde{\theta}_l(\mathbf{r})$ and integrating over the spatial coordinates reformulates the OEP equation as a matrix equation:

$$\sum_{k}^{n_{\mathrm{aux}}} \rho_k^{\mathrm{s}} A_{kl} = b_l \Longleftrightarrow \sum_{k}^{n_{\mathrm{aux}}} \left(A^{-1}\right)_{kl} b_l = \rho_k^{\mathrm{s}}, \tag{4.8}$$

where the matrix form of the response function $A_{kl} = \left\langle \tilde{\theta}_k \middle| \chi \middle| \tilde{\theta}_l \right\rangle$ and the right-hand side vector $b_l = \left\langle b \middle| \tilde{\theta}_l \right\rangle$.

The exact form of b_l is dependent on the problem that is being solved. In contrast, for a given finite orbital basis set containing n_{orb} basis functions, the matrix elements A_{kl} always take the same form (where we have assumed that the orbitals ϕ_i, ϕ_a are real):

$$A_{kl} = 2 \sum_{\sigma} \sum_{i=1}^{N_\sigma} \sum_{a=N_\sigma+1}^{n_{\text{orb}}} \frac{\langle \phi_i | \, \tilde{\theta}_k \, | \phi_a \rangle \, \langle \phi_i | \, \tilde{\theta}_l \, | \phi_a \rangle}{\epsilon_i - \epsilon_a}. \tag{4.9}$$

With a given form for b_l, the above matrix OEP equation could be solved numerically. Before directly trying this, however, we must remember that the solution to the OEP equation for the effective potential is only determined up to a constant. Therefore, a procedure must be implemented that is able to fix this constant, preferably to a physically meaningful quantity.

4.1.3 The constrained minimisation of density functionals

To fix the screening charge the constrained minimisation approach [100, 101] will be used, which begins by employing the standard total energy functional in DFT, where a density functional approximation is used to model the exchange-correlation functional, $E_{\text{xc}}^{\text{DFA}}[\rho]$:

$$E^{\text{DFA}}[\rho] = T_{\text{s}}[\rho] + U[\rho] + \int \mathrm{d}\mathbf{r} \; v_{\text{en}}(\mathbf{r}) \rho(\mathbf{r}) + E_{\text{xc}}^{\text{DFA}}[\rho], \tag{4.10}$$

where $T_{\text{s}}[\rho]$ is the non-interacting kinetic energy functional, $U[\rho]$ is the Hartree energy functional and $v_{\text{en}}(\mathbf{r})$ is the external potential.

Following the OEP method, the KS orbitals are constructed to satisfy their own set of single-particle equations, experiencing an effective potential $v_{\text{eff}}(\mathbf{r})$:

$$\left[-\frac{\nabla^2}{2} + v_{\text{en}}(\mathbf{r}) + v_{\text{eff}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \tag{4.11}$$

The total energy functional is now minimised with two constraints applied to the effective potential $v_{\text{eff}}(\mathbf{r})$. In this form of the OEP equation, the effective potential

takes the place of the Hartree-exchange-correlation potential $v_{\mathrm{Hxc}}(\mathbf{r})$. However, due to the constraints imposed on the effective potential detailed below, in general $v_{\mathrm{eff}}(\mathbf{r}) \neq v_{\mathrm{Hxc}}(\mathbf{r})$.

It is also possible to factor out the Hartree potential $v_{\mathrm{H}}(\mathbf{r})$, such that the KS orbitals satisfy the following KS equations:

$$\left[-\frac{\nabla^2}{2} + v_{\mathrm{en}}(\mathbf{r}) + v_{\mathrm{H}}(\mathbf{r}) + v_{\mathrm{eff}}(\mathbf{r})\right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \tag{4.12}$$

where in this form, the effective potential mimics the exchange-correlation potential $v_{\mathrm{xc}}(\mathbf{r})$. However, attention will firstly be focused on the formulation where the Hartree potential is included in the effective potential in Equation 4.11.

As discussed previously, the integral of the screening density is a measure of the strength of self-interactions present in a given density functional approximation. As such, the screening density is treated as the fundamental quantity within the constrained method. It is already established that analytically the exact self-interaction free xc potential should have a Hxc screening charge equal to $N - 1$. Therefore, the first constraint that we impose on the effective potential is that:

$$Q_{\mathrm{scr}} = \int \mathrm{d}\mathbf{r}\ \rho_{\mathrm{scr}}(\mathbf{r}) = N - 1. \tag{4.13}$$

Most of the commonly used density functional approximations do not satisfy this constraint out of the box, and instead show an effective screening charge of N, i.e. are fully self-interaction contaminated, as this implies that each electron experiences an effective repulsion from N electrons (thus including itself).

It is important to note here that the violation of the constraint 4.13 by a density functional approximation does not imply that the well-known sum rule for the exchange-correlation hole, $\int \mathrm{d}\mathbf{r}\ \rho_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}') = -1$[111], is also violated. Although the quantities $\rho_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$ and $\rho^{\mathrm{xc}}_{\mathrm{scr}}(\mathbf{r})$ appear similar, they are not directly related quantities and should not be confused with one another.

The second constraint that we impose on the screening density is that it must be

strictly positive across all space:

$$\rho_{\mathrm{scr}}(\mathbf{r}) \geq 0. \tag{4.14}$$

Although the first constraint is based on a analytical property of the exact KS potential, the second constraint imposed on the screening density has two differing motivations and justifications.

Firstly, it is somewhat intuitive to assume that the Hxc screening density is positive, as it is effectively the repulsive potential experienced by each electron from the other $N-1$ electrons. This is not however an exact property satisfied by the exact KS potential and the presence of negative regions of screening density within the vicinity of a system are not strictly prohibited. It is worth noting however that although it is not an exact property, previous studies that employed this constraint demonstrated that it yields good results [100, 101, 112].

Secondly, the inclusion of the positivity constraint is what allows the mathematical problem of constrained minimisation to be well-posed in the complete orbital and auxiliary basis set limits. If we do not include this constraint, the screening density has a tendency to decompose into two components, one with the natural screening charge of the underlying density functional approximation, $Q_{\mathrm{scr}}^{\mathrm{DFA}} = N$, and the other with $Q_{\mathrm{scr}}^{\mathrm{SIC}} = -1$ (where SIC stands for self-interaction correction), that arises from the overall $N-1$ constraint on the total screening density. The first component stemming from the approximation employed will always remain in the vicinity of the molecule, as this region is energetically important. The second component however will be pushed out to infinity as the size of the auxiliary basis set is increased, as this is energetically the most favourable solution. Remember, the underlying functional naturally wants to have an overall screening charge of N and this method attempts to reduce this to $N-1$. With the positivity constraint however this behaviour is not possible, and the system is forced to cope with this unnatural screening charge, by incurring a slight penalty to the total energy.

Due to the constraints on the effective potential, the total energy resulting will

always be slightly higher than the unconstrained solution. However, in the results section it is shown that the total energy difference between the two methods is negligible (10^{-4} eV).

To perform the constrained minimisation under these two described constraints we reformulate the total energy expression to be minimised as an objective functional, $G[\rho_{\mathrm{scr}}]$ to be minimised:

$$G[\rho_{\mathrm{scr}}] = E^{\mathrm{DFA}}[\rho_{\mathrm{scr}}] + E^{\mathrm{SIC}}[\rho_{\mathrm{scr}}] + E^{\mathrm{POS}}[\rho_{\mathrm{scr}}], \tag{4.15}$$

where $E^{\mathrm{DFA}}[\rho_{\mathrm{scr}}]$ is total energy functional defined in Equation 4.10 and $E^{\mathrm{SIC}}[\rho_{\mathrm{scr}}]$ is the term that enforces the $N-1$ constraint of the screening density:

$$E^{\mathrm{SIC}}[\rho_{\mathrm{scr}}] = -\lambda \left[\int \mathrm{d}\mathbf{r}\ \rho_{\mathrm{scr}}(\mathbf{r}) - (N-1) \right], \tag{4.16}$$

and $E^{\mathrm{POS}}[\rho_{\mathrm{scr}}]$ is the penalty term that imposes the positivity constraint:

$$E^{\mathrm{POS}}[\rho_{\mathrm{scr}}] = \eta \left[\int \mathrm{d}\mathbf{r}\ |\rho_{\mathrm{scr}}(\mathbf{r})| - (N-1) \right];\ \eta >> 1. \tag{4.17}$$

At its minimum the functional derivative of $G[\rho_{\mathrm{scr}}]$ with respect to the screening density must vanish, i.e.:

$$\frac{\delta G[\rho_{\mathrm{scr}}]}{\delta \rho_{\mathrm{scr}}(\mathbf{x})} = 0. \tag{4.18}$$

Evaluating the functional derivatives for the three components of the objective functional we find:

$$\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\ \frac{\delta E^{\mathrm{DFA}}[\rho_{\mathrm{scr}}]}{\delta \rho(\mathbf{r})} \frac{\delta \rho(\mathbf{r})}{\delta v_{\mathrm{eff}}(\mathbf{r}')} \frac{\delta v_{\mathrm{eff}}(\mathbf{r}')}{\delta \rho_{\mathrm{scr}}(\mathbf{x})} - \lambda + \eta\, \mathrm{sgn}\left(\rho_{\mathrm{scr}}(\mathbf{x})\right) = 0, \tag{4.19}$$

where the functional chain rule is used to evaluate $\delta E^{\mathrm{DFA}}[\rho_{\mathrm{scr}}]/\delta \rho_{\mathrm{scr}}(\mathbf{x})$.

The first derivative can be calculated by making use of Equation 3.21 and the definitions of the Hartree and exchange-correlation potentials:

$$\frac{\delta E^{\mathrm{DFA}}[\rho_{\mathrm{scr}}]}{\delta \rho(\mathbf{r})} = v_{\mathrm{Hxc}}^{\mathrm{DFA}}(\mathbf{r}) - v_{\mathrm{eff}}(\mathbf{r}), \tag{4.20}$$

where:

$$v_{\mathrm{eff}}(\mathbf{r}) = -\frac{\delta T_{\mathrm{s}}[\rho]}{\delta \rho(\mathbf{r})}. \tag{4.21}$$

Using the definition of the response function 3.57 we know that:

$$\frac{\delta\rho(\mathbf{r})}{\delta v_{\text{eff}}(\mathbf{r}')} = \chi(\mathbf{r},\mathbf{r}'). \tag{4.22}$$

It is also easy to evaluate the third derivative using the definition of the effective potential 4.2:

$$\frac{\delta v_{\text{eff}}(\mathbf{r}')}{\delta\rho_{\text{scr}}(\mathbf{x})} = \frac{1}{|\mathbf{r}'-\mathbf{x}|}. \tag{4.23}$$

Putting everything together the OEP equation is given by:

$$\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \left[v_{\text{Hxc}}^{\text{DFA}}(\mathbf{r}) - v_{\text{eff}}(\mathbf{r})\right] \frac{\chi(\mathbf{r},\mathbf{r}')}{|\mathbf{r}'-\mathbf{x}|} - \lambda + \eta\,\text{sgn}\left(\rho_{\text{scr}}(\mathbf{x})\right) = 0. \tag{4.24}$$

Rearranging this equation and inserting the expansion of the effective potential in terms of the screening density we arrive at a form of the OEP equation for the screening density:

$$\int \mathrm{d}\mathbf{y}\ \tilde{\chi}(\mathbf{x},\mathbf{y})\rho_{\text{scr}}(\mathbf{y}) = b(\mathbf{x}) - \lambda + \eta\,\text{sgn}\left(\rho_{\text{scr}}(\mathbf{x})\right), \tag{4.25}$$

where:

$$\tilde{\chi}(\mathbf{x},\mathbf{y}) = \iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \frac{\chi(\mathbf{r},\mathbf{r}')}{|\mathbf{r}'-\mathbf{x}||\mathbf{r}-\mathbf{y}|}, \tag{4.26}$$

$$b(\mathbf{x}) = \iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}' \frac{\chi(\mathbf{r},\mathbf{r}')}{|\mathbf{r}'-\mathbf{x}|} v_{\text{Hxc}}^{\text{DFA}}(\mathbf{r}). \tag{4.27}$$

The next step is to transform this equation into the finite basis set representation as demonstrated in § 4.1.2. Inserting the expansion 4.6 result in a modified version of Equation 4.8 which incorporates the two constraints:

$$\sum_{k}^{n_{\text{aux}}} \rho_k^{\text{s}} A_{kl} = b_l - \lambda X_l + \eta Y_l, \tag{4.28}$$

or alternatively:

$$\sum_{k}^{n_{\text{aux}}} \left(A^{-1}\right)_{kl} \left[b_l - \lambda X_l + \eta Y_l\right] = \rho_k^{\text{s}}, \tag{4.29}$$

where A_{kl} has the same definition as in Equation 4.9, $b_l = \left\langle b \middle| \tilde{\theta}_l \right\rangle$, and:

$$X_l = \int \mathrm{d}\mathbf{r}\ \theta_l(\mathbf{r}); \quad Y_l = \int \mathrm{d}\mathbf{r}\ \theta_l(\mathbf{r})\,\text{sgn}\left(\rho_{\text{scr}}(\mathbf{x})\right). \tag{4.30}$$

The value of the Lagrange multiplier, λ, can be found by expanding the screening charge in the auxiliary basis set and substituting Equation 4.29 for the coefficients, yielding:

$$\lambda = \frac{\sum_{k,l}^{n_{\text{aux}}} X_k \left(A^{-1}\right)_{kl} \left[b_l + \eta Y_l\right] - Q_{\text{scr}}}{\sum_{k,l}^{n_{\text{aux}}} X_k \left(A^{-1}\right)_{kl} X_l}. \tag{4.31}$$

In this representation, the matrix A_{kl} contains extremely small eigenvalues. As such, when attempting to construct its inverse a singular value decomposition (SVD) is required to remove the eigenvectors with near-zero eigenvalues. Even with these components removed however, the effective potential still can exhibit undesirable oscillations [113]. One could make use of the KLI or CEDA approximations to find robust solutions for the OEP equation, though both of these suffer from some accuracy loss (although we shall see that CEDA performs well for the systems studied in this thesis). Therefore, a form of the OEP equations for finite basis sets needs to be constructed that can systematically avoid these numerical issues.

4.1.4 The complement of the response function

In principle, the above equations could be used to solve the OEP equation. However, even in this form, the OEP equation suffers from several numerical issues that must be addressed in order to find robust solutions. The most obvious example of these problems is the presence of extremely unphysical oscillations in the effective potential.

Over the years there have been numerous approaches put forward with the aim of finding robust solutions to the OEP equation which avoid these issues. One commonly used approach is to construct the orbital and auxiliary basis sets in such a way that the orbital basis is always well-converged with respect to the auxiliary basis [70, 71]. This is achieved by selecting an auxiliary basis and than gradually increasing the size or of the orbital basis until convergence is achieved. Although this approach has proved successful, it is very computationally expensive.

Instead the 'complement' method, introduced by Gidopoulos and Lathiotakis [113], can be used to construct density-density response functions that avoid these issues. This method has been tested extensively, and it has been shown to systematically alleviate the difficulties faced in solving the OEP equation for closed-shell systems. In the following chapters, this method will also be extended to construct response functions for open shell systems and for ensemble with varying numbers of electrons.

Both sides of the OEP equation depend on the response function $\chi(\mathbf{r},\mathbf{r}')$. As such, finding an accurate matrix representation, A_{kl}, is crucial to solving the OEP equation. To discuss the complement method, it is important to revisit the general form of the response function (from Equation 3.59):

$$\chi(\mathbf{r},\mathbf{r}') = 2\sum_{\sigma}^{\uparrow\downarrow}\sum_{i=1}^{N^\sigma}\sum_{a=N^\sigma+1}^{\infty}\frac{\phi_i(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a(\mathbf{r}')}{\epsilon_i-\epsilon_a} = \chi^{\uparrow}(\mathbf{r},\mathbf{r}')+\chi^{\downarrow}(\mathbf{r},\mathbf{r}'), \quad (4.32)$$

where it is assumed that the KS orbitals are real (which is the case for finite systems with zero external applied magnetic field).

The infinite sum over all of the unoccupied orbitals present in this form is clearly something we want avoid if possible in a finite basis set implementation. Naively is would be possible to truncate this summation at the number of orbital basis set functions, n_{orb}, to simplify its construction. Unfortunately however, this leads to a poor representation for $\chi(\mathbf{r},\mathbf{r}')$, as the higher unoccupied orbital terms in the sum have a significant influence on the response function. Instead, the summation is split over all unoccupied orbitals into two parts:

$$\chi^{\lambda}(\mathbf{r},\mathbf{r}') = \chi_0(\mathbf{r},\mathbf{r}')+\lambda\bar{\chi}(\mathbf{r},\mathbf{r}'), \quad (4.33)$$

where for $\lambda = 1$, $\chi_0(\mathbf{r},\mathbf{r}')$ can be expressed in terms of the known n_{orb} lowest lying orbitals:

$$\chi_0(\mathbf{r},\mathbf{r}') = 2\sum_{\sigma}^{\uparrow\downarrow}\sum_{i=1}^{N_\sigma}\sum_{a=N_\sigma+1}^{n_{\text{orb}}}\frac{\phi_i(\mathbf{r})\phi_i(\mathbf{r}')\phi_a(\mathbf{r})\phi_a(\mathbf{r}')}{\epsilon_i-\epsilon_a} = \chi_0^{\uparrow}(\mathbf{r},\mathbf{r}')+\chi_0^{\downarrow}(\mathbf{r},\mathbf{r}'), \quad (4.34)$$

and $\bar{\chi}(\mathbf{r},\mathbf{r}')$ is termed the 'complement' of the response function, constructed from the rest of the unoccupied orbitals that do not lie within the orbital basis set:

$$\bar{\chi}(\mathbf{r},\mathbf{r}') = 2\sum_{\sigma}^{\uparrow\downarrow}\sum_{i=1}^{N_\sigma}\sum_{a=n_{\text{orb}}+1}^{\infty}\frac{\phi_i(\mathbf{r})\phi_i(\mathbf{r}')\phi_a(\mathbf{r})\phi_a(\mathbf{r}')}{\epsilon_i-\epsilon_a} = \bar{\chi}^{\uparrow}(\mathbf{r},\mathbf{r}')+\bar{\chi}^{\downarrow}(\mathbf{r},\mathbf{r}'). \quad (4.35)$$

In all other OEP methods, the complement part of the response function is typically ignored and one must resort to very large basis sets to ensure an accurate representation of $\chi_0(\mathbf{r},\mathbf{r}')$ is constructed. Doing so clearly has a large computational impact, as larger basis sets leads to longer computation times.

In this form however, the complement still contains the difficult sum that we wish to avoid. This sum can be bypassed by first employing the Unsöld approximation [76] to simplify the denominator:

$$\bar{\chi}(\mathbf{r},\mathbf{r}') \approx -\frac{2}{\Delta}\sum_{\sigma}^{\uparrow\downarrow}\sum_{i=1}^{N_\sigma}\sum_{a=n_{\mathrm{orb}}+1}^{\infty}\phi_i(\mathbf{r})\phi_i(\mathbf{r}')\phi_a(\mathbf{r})\phi_a(\mathbf{r}'), \tag{4.36}$$

where the constant $-\Delta$ is the common energy denominator that replaces the different $\epsilon_i - \epsilon_a$.

The insertion of the completeness relation 3.69 into the sum over the unoccupied orbitals then yields:

$$\bar{\chi}(\mathbf{r},\mathbf{r}') \approx -\frac{2}{\Delta}\sum_{\sigma}^{\uparrow\downarrow}\sum_{i=1}^{N_\sigma}\phi_i(\mathbf{r})\phi_i(\mathbf{r}')\left[\delta(\mathbf{r}-\mathbf{r}') - \sum_{k=1}^{n_{\mathrm{orb}}}\phi_k(\mathbf{r})\phi_k(\mathbf{r}')\right]. \tag{4.37}$$

The final term in the above equation can then be split into two parts, the first containing all of the occupied terms, and the other the lowest unoccupied orbitals up to n_{orb}:

$$\sum_{k=1}^{n_{\mathrm{orb}}}\phi_k(\mathbf{r})\phi_k(\mathbf{r}') = \sum_{j=1}^{N_\sigma}\phi_j(\mathbf{r})\phi_j(\mathbf{r}') + \sum_{a=N_\sigma+1}^{n_{\mathrm{orb}}}\phi_a(\mathbf{r})\phi_a(\mathbf{r}'). \tag{4.38}$$

On the right-hand side, the second sum, which is also present in $\chi^0(\mathbf{r},\mathbf{r}')$, can be safely discarded, as in the desired limit of $\lambda \to 0$ its contribution to the total response function vanishes.

By applying the same procedure to the construction of the right-hand side of the OEP equation, $b(\mathbf{r})$, the splitting of the response function results in:

$$\begin{aligned} b^0(\mathbf{r}) &= 2\sum_{\sigma}^{\uparrow\downarrow}\sum_{i=1}^{N_\sigma}\sum_{a=N_\sigma+1}^{n_{\mathrm{orb}}}\frac{\phi_i(\mathbf{r})\phi_a(\mathbf{r})}{\epsilon_i-\epsilon_a}\int \mathrm{d}\mathbf{r}'\phi_i(\mathbf{r}')\left[v_{\mathrm{H}}(\mathbf{r}') + v_{\mathrm{xc}}^{\mathrm{DFA}}(\mathbf{r}')\right]\phi_a(\mathbf{r}') \\ &= b^{0\uparrow}(\mathbf{r}) + b^{0\downarrow}(\mathbf{r}), \end{aligned} \tag{4.39}$$

and:

$$\begin{aligned}\bar{b}(\mathbf{r}) &= -\frac{2}{\Delta}\rho(\mathbf{r})\left[v_{\mathrm{H}}(\mathbf{r})+v_{\mathrm{xc}}^{\mathrm{DFA}}(\mathbf{r})\right] \\ &\quad -\sum_{k=1}^{n_{\mathrm{orb}}}\phi_k(\mathbf{r})\int \mathrm{d}\mathbf{r}'\ \phi_i(\mathbf{r}')\left[v_{\mathrm{H}}(\mathbf{r}')+v_{\mathrm{xc}}^{\mathrm{DFA}}(\mathbf{r}')\right]\phi_k(\mathbf{r}')\Bigg\} \\ &= \bar{b}^{\uparrow}(\mathbf{r})+\bar{b}^{\downarrow}(\mathbf{r}),\end{aligned} \tag{4.40}$$

where the total $b(\mathbf{r})$ is given the the sum of the two terms, i.e. $b^{\lambda}(\mathbf{r}) = b^{0}(\mathbf{r})+\lambda\bar{b}(\mathbf{r})$.

Finally, the OEP equation can be written as:

$$\int \mathrm{d}\mathbf{r}'\left[\chi^{0}(\mathbf{r},\mathbf{r}')+\lambda\bar{\chi}(\mathbf{r},\mathbf{r}')\right]v_{\mathrm{eff}}^{\lambda}(\mathbf{r}') = b^{0}(\mathbf{r})+\lambda\bar{b}(\mathbf{r}). \tag{4.41}$$

A similar methodology is used to convert this form of the OEP equation into a matrix representation as we previously detailed. As a result, the response function matrix is given by:

$$A_{kl} = A_{kl}^{0}+\alpha\bar{A}_{kl}, \tag{4.42}$$

where $\alpha = \frac{\lambda}{\Delta}$, A_{kl}^{0} is the same as in Equation 4.9, and the complement term $\bar{A}_{kl}$ is given by:

$$\bar{A}_{kl} = -\sum_{\sigma}\sum_{i=1}^{N_\sigma}\left\{\langle\phi_i|\,\tilde{\theta}_k\tilde{\theta}_l\,|\phi_i\rangle-\sum_{j=1}^{N_\sigma}\langle\phi_i|\,\tilde{\theta}_k\,|\phi_j\rangle\,\langle\phi_i|\,\tilde{\theta}_l\,|\phi_j\rangle\right\}. \tag{4.43}$$

In Ref. [113], the OEP equation is shown to contain a discontinuity at $\lambda = 0$. This difference between the effective potentials $v^{\lambda=0}$ and $v^{\lambda\to 0^{+}}$, implies that for a finite orbital basis representation the effective potential does not (in general) coincide with the limit of the complete orbital basis result for $\lambda \to 0$, $\lambda > 0$. Only if the orbital basis used is large enough to overlap or cover completely the space of the auxiliary basis would these two potential converge smoothly. The inclusion of the complement for small $\lambda > 0$ (and thus small $\alpha > 0$) solves the issue of unphysical oscillations in the potential; though it is required to trial several values to find the smallest value of α which remains (but slightly larger than the near-zero eigenvalues of χ^{0}) that completes the space of the response function. An interesting point to note is that in the limit $\lambda \to \infty$, the response function is dominated by the complement contribution, $\bar{\chi}(\mathbf{r},\mathbf{r}')$, and reduces to the same representation

for $\chi(\mathbf{r}, \mathbf{r}')$ as in the CEDA approximation. So, we now have a meaningful OEP equation in the whole range of $0 < \alpha < \infty$ with a response function that varies smoothly and is always invertible up to a constant.

Since Ref. [113] was published, I have carried out extensive testing and refinement of this method for solving the OEP equation, with a new, more efficient and flexible routine written from the ground up. Notably, I streamlined the construction of the right-hand side of the OEP equation, $b(\mathbf{r})$, to improve computational efficiency for all the functionals implemented.

I mainly worked on an original study investigating the possibility of relaxing the positivity constraint to test if it is required, as this constraint proves to be a computational bottleneck (having to integrate the screening charge over the real-space grid to check to negative regions is extremely slow).

All of the results presented in the following section were obtained using the newly updated version of the HIPPO code; building on the original implementations of Gidopoulos, Lathiotakis, and Pitts [100, 112].

4.2 Results

4.2.1 The constrained xc potential

In the previous section, it was established that one of the aims of the constrained method was to correct the asymptotic decay of the xc potential. As the exact xc potential decays as $\frac{-1}{r}$, enforcing the xc-screening density to integrate to -1 is expected to systematically embed this property of the exact xc potential into any calculated effective potential. In Figure 4.1 the difference between the unconstrained and constrained xc potentials is very apparent - the constrained calculations is shifted downwards from the unconstrained solution, and exhibits a different decay as $r \to \infty$.

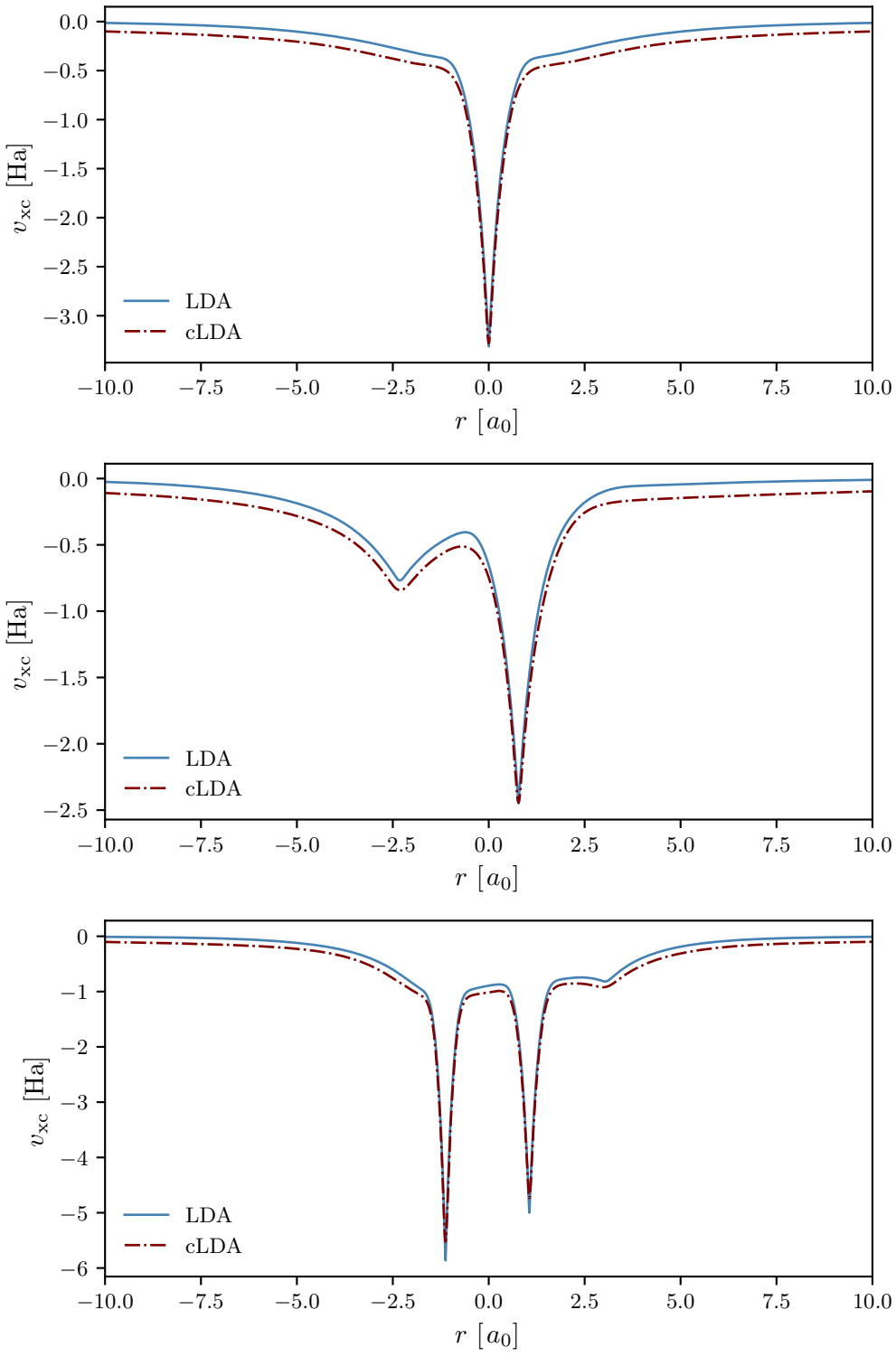


Figure 4.1: The unconstrained and constrained xc potentials calculated for (top) the Beryllium atom (Be), (middle) Lithium hydride (LiH), and (bottom) Hydrogen isocyanide (HCN), using a cc-pVTZ orbital basis and an uncontracted cc-pVDZ auxiliary basis.

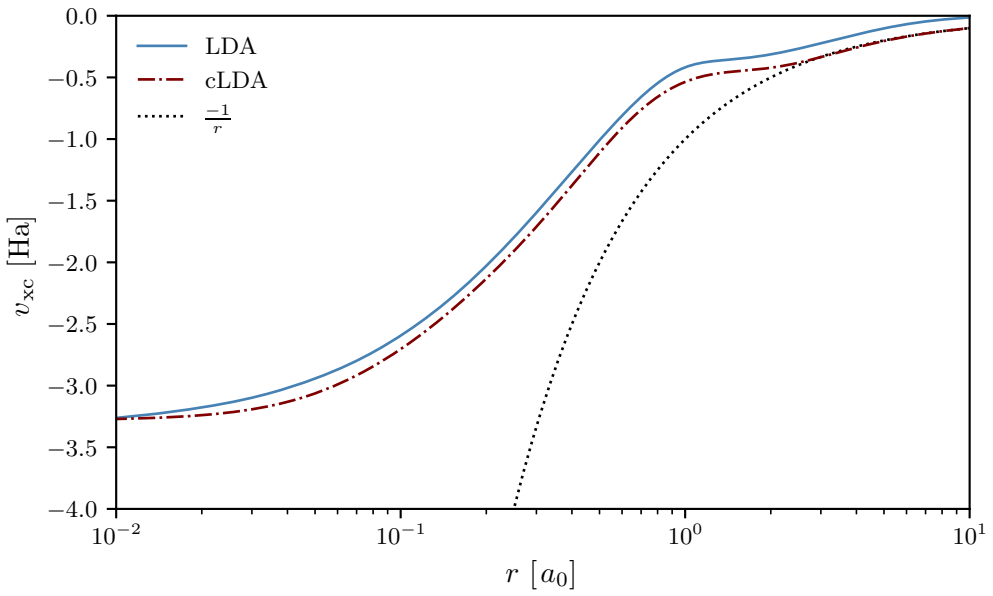


Figure 4.2: A comparison of the asymptotic decay of the xc potential obtained from unconstrained and constrained LDA for a Beryllium atom.

The asymptotic decay can be investigated further by plotting the xc potentials on a logarithmic scale in the large r region. Figure 4.2 plots both the unconstrained LDA potential for the Beryllium atom, compared to the constrained LDA potential to compare their asymptotic decay. It is clear that the cLDA potential follows the exact $\frac{-1}{r}$ decay expected, whereas the unconstrained LDA curve deviates from the exact decay, and this behaviour is observed in all calculated $N-1$ constrained effective potentials. All of these above calculations were performed with the positivity constraint enforced on the effective potential, however, in the following sections, we will instead present results obtained when this constraint is relaxed.

4.2.2 Lifting the constraint of positivity

The positivity constraint was originally enforced on the screening density to ensure that in the complete orbital and auxiliary basis set limits, the tendency for the screening density to decompose into two distinct components, one with the natural screening charge of the underlying DFA and the other with a screening charge of

−1. In these large basis set limits, the negative component (the self-interaction correction) has the tendency to move away from the system as it becomes the more energetically favourable solution. Enforcing the positivity constraint on the total screening density renders it impossible for this leaking of negative change far from the system to take place, allowing the constrained minimisation to be mathematically well-posed (see § 4.1.3 for a more detailed discussion).

However, checking and enforcing this constraint over all space is time-consuming and computationally costly. Furthermore, as the screening density is a fictitious density, is there any reason why regions of this density, specifically in the region of the molecule could be negative in some regions within the vicinity of the system? If it were possible to relax this restriction in practice without incurring any undesired behaviour, it would not only dramatically reduce the computation times for constrained calculations but it would also allow for the possibility of modelled systems with regions of negative effective screening density.

In this section the effects of relaxing this constraint and relying solely on the constraint of $Q_{\text{scr}} = N - 1$ are discussed. Each of the following calculations presented were performed using a finite α value of 10^{-2} across a range of orbital and auxiliary basis sets. This was selected by performing successive calculations with a smaller value until the effective potential does not change. Although this selected value for α is in principle system dependent, for the four systems presented here convergence in the potentials was obtained at roughly the same value of $\alpha = 10^{-2}$.

As the positivity constraint allowed the solution of the OEP equation to remain well-posed in the infinite basis set limit, we expect to see a change in the screening density as the auxiliary basis set size is increased. If however this change is not significant, then it can be concluded that relaxing the positivity constraint is indeed a safe thing to do. To investigate the removal of the positivity constraint, the convergence of the xc potential and screening density of a small selection of systems are analysed.

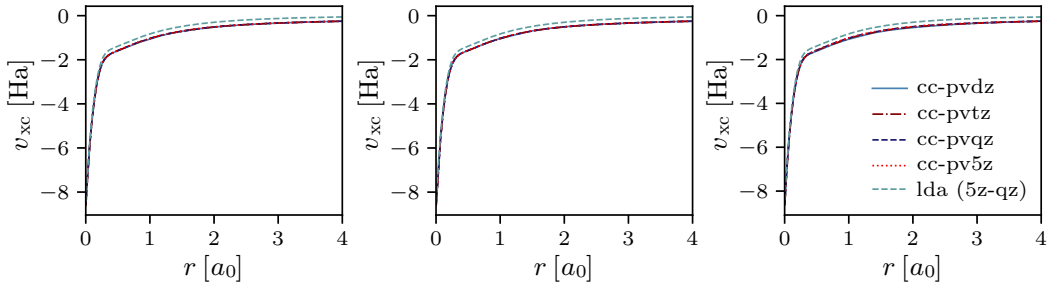


Figure 4.3: The effective potential calculated for a Neon atom for range of basis set combinations. The legend denotes the size of the orbital basis set used for each calculation and is consistent across all three subplots. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively.

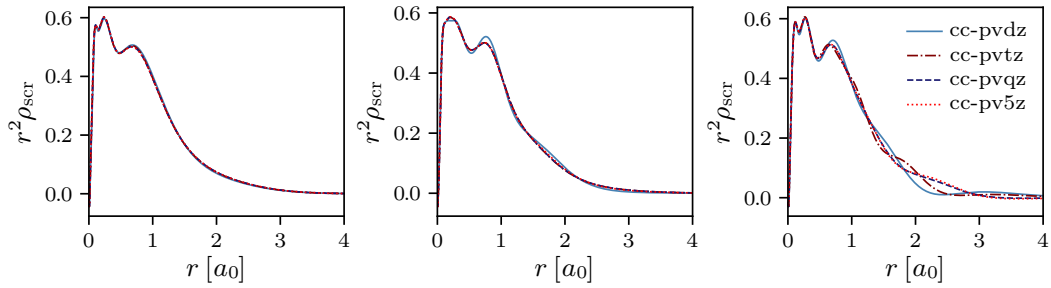


Figure 4.4: The screening density plotted as $r^2\rho_{\mathrm{scr}}(r)$ for a Neon atom for range of basis set combinations. The legend denotes the size of the orbital basis set used for each calculation and is consistent across all three subplots. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively.

Figure 4.3 shows the convergence of the cLDA xc potential for the Neon atom with respect to the size of the orbital basis set for three different auxiliary basis sets; each of these calculations was performed without the positivity constraint. The constrained potentials are also compared to the unconstrained LDA potential obtained with a cc-pV5Z orbital basis set (dashed green line). It is clear from these plots that the variations in the combinations of basis sets has very little effect on the OEP potential, provided the calculation can converge. We can also inspect the screening densities calculated for the same basis set combinations, which are presented in Figure 4.4. As the screening density is very hard to converge this is

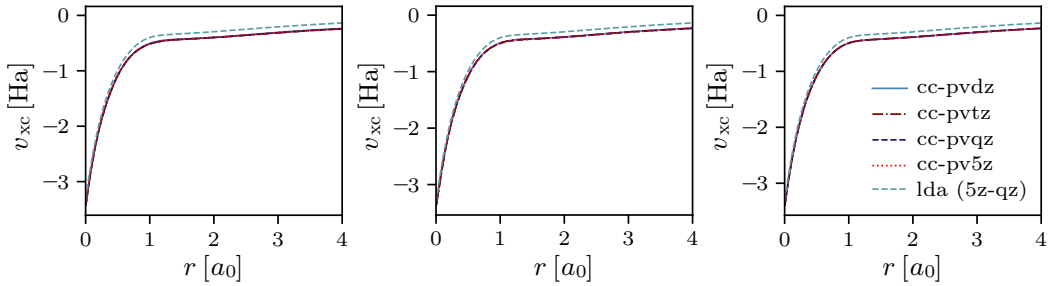


Figure 4.5: The effective potential calculated for a Beryllium atom for range of basis set combinations. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively, with the orbital basis indicated by the different coloured curves.

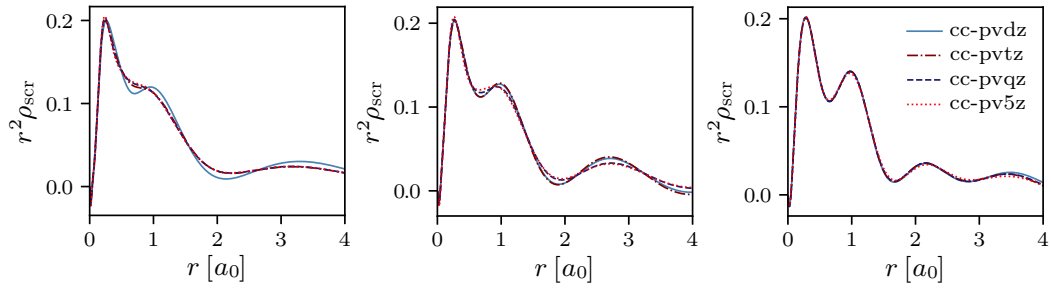


Figure 4.6: The screening density plotted as $r^2\rho_{\text{scr}}(r)$ for a Beryllium atom for range of basis set combinations. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively, with the orbital basis indicated by the different coloured curves.

a very strict test for overall convergence. When plotting the screening density for an atom, it makes sense to plot $r^2\rho_{\text{scr}}(r)$ so it is easier to observe the decay of the screening density; otherwise the behaviour near the atom will dominate. For Neon, the convergence of the xc potential is very good despite not enforcing the positivity constraint. There is also no sign of any negative screening density moving away from the atom, even for the very large uncontracted cc-pVQZ auxiliary basis.

A similar picture can be seen when comparing the results from Neon above to those of an atom of Beryllium. In Figure 4.5 the effective potential is extremely well converged even at modest orbital basis set size. The screening charge plots in Figure 4.6 also appear to be well converged, despite the presence of some small

oscillations in the uncontracted cc-pVQZ results.

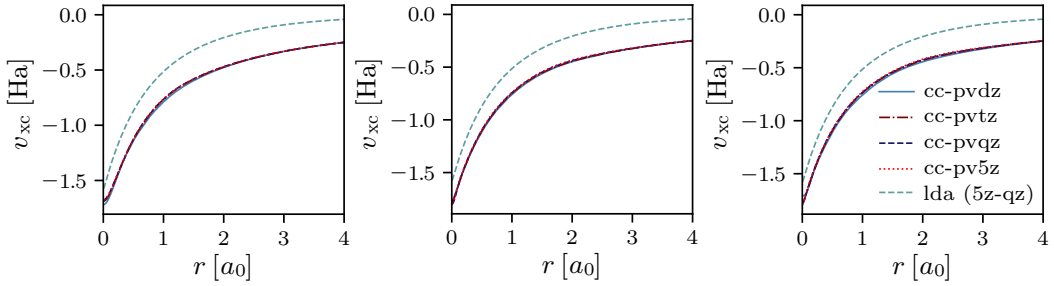


Figure 4.7: The effective potential calculated for a Helium atom for range of basis set combinations. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively, with the orbital basis indicated by the different coloured curves.

If the system is made even smaller however, more severe issues begin to manifest. As for Neon and Beryllium, Figure 4.7 shows good convergence in the effective potential across all basis set combinations for the Helium atom. In the screening density plots in Figure 4.8 however, we note that convergence for a fixed auxiliary basis and increasing orbital basis set size is slower. Also unlike in the previous two systems discussed, for Helium there is a region of negative screening density that appears to shift away from the atom as the size of the auxiliary basis set increases. The presence of a negative region of screening density, although perhaps unintuitive, is not necessarily a sign of bad convergence; the movement of this neg-

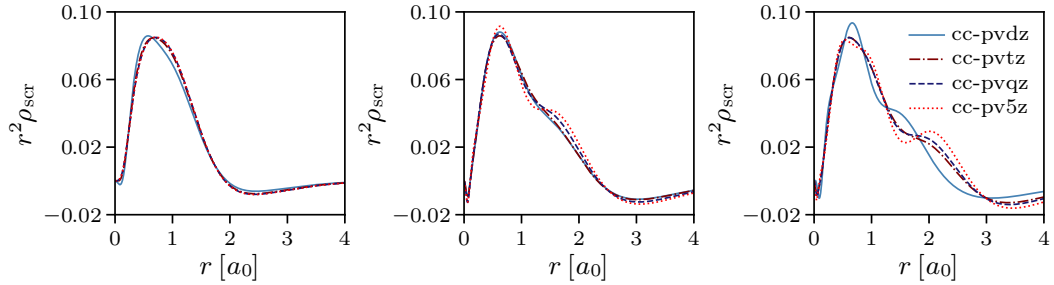


Figure 4.8: The screening density plotted as $r^2\rho_{\mathrm{scr}}(r)$ for a Helium atom for range of basis set combinations. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively, with the orbital basis indicated by the different coloured curves.

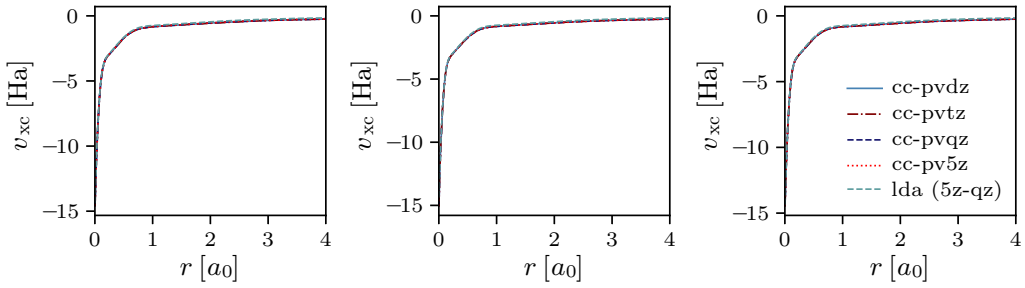


Figure 4.9: The effective potential calculated for a Chloride anion for range of basis set combinations. The legend denotes the size of the orbital basis set used for each calculation and is consistent across all three subplots. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively.

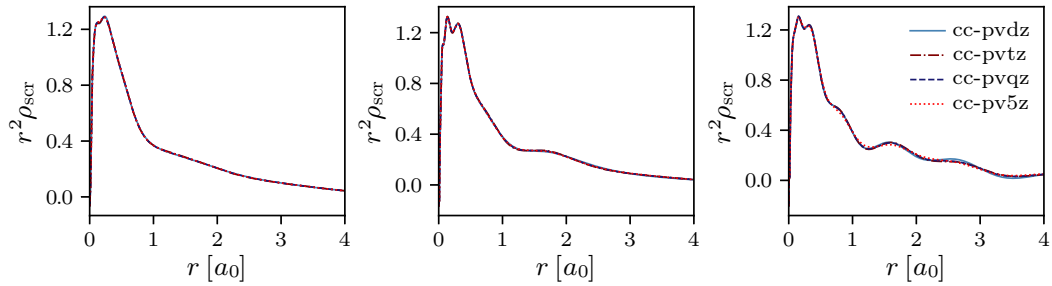


Figure 4.10: The screening density plotted as $r^2\rho_{\text{scr}}(r)$ for a Chloride anion for range of basis set combinations. The legend denotes the size of the orbital basis set used for each calculation and is consistent across all three subplots. Each of the subplots has a fixed uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, (right) cc-pVQZ auxiliary basis set respectively.

ative density away from the atom with increasing auxiliary basis set size is however an undesirable outcome. Without the positivity constraint to force constrained minimisation to remain well-posed, the minimum energy solution is obtained by shifting the negative region away from the system in an attempt to converge to a solution with a screening charge of N in the neighbourhood of the atom. This effect is far more prevalent in Helium due to the relative magnitude of difference between N and $N-1$. As the number of electrons increases, the relative difference between these two values decreases. Therefore, in Helium we expect this effect to be strongest; which is indeed confirmed in the results presented.

The final system to discuss is the Chloride anion (Cl^-). Negatively charged ions are notoriously difficult for LDA and regular, unconstrained calculations, are often predicted to be unbound. This result was already observed in Table 4.4, where the calculated HOMO eigenvalue for the Cl^- ion was 2.731 eV. The constrained method however correctly predicts these systems as bound ions (again see Table 4.4), and it is clear from Figs. 4.9 and 4.10 that convergence for the Chloride anion is excellent across all basis sets tested.

Through discussing the screening densities and OEP potentials obtained for these systems, it is clear that at very large auxiliary basis sets and a small number of electrons the unphysical tendency for the screening charge hole to separate into two components, with the self-interaction term (the negative density) being pushed away from the molecule, does indeed manifest. It is however possible to conclude that for moderately sized auxiliary basis sets and systems with more than a few electrons the positivity constraint can be relaxed without incurring any severe issues. Lifting the positivity constraint allows a major computational bottleneck to be bypassed and allows for the flexibility of having negative regions of screening densities in the neighbourhood of a system.

4.2.3 Ground state energies

It is important that the constrained OEP method can yield ground state total energies that are near to the unconstrained solutions. Placing additional constraints on the system through the screening charge and positivity constraints, it is to be expected that total energies will be slightly higher than a regular DFT calculation.

In Table 4.1 ground state total energy results obtained from both the constrained method and standard DFT for the LDA functional for a range of small closed shell systems. These results were obtained without the constraint of positivity enforced on the screening density, however (with the exception of the smallest systems, He, Be, and LiH), total energies calculated with this constraint enforced are very sim-

System	E^{LDA} (Ha)	E^{cLDA} (Ha)	ΔE (mHa)
He	-2.8717	-2.8716	0.1209
Be	-14.5202	-14.5201	0.0273
C	-37.5140	-37.5136	0.3611
O	-74.5801	-74.5799	0.1979
Ne	-128.4218	-128.4218	0.0297
Mg	-199.3689	-199.3686	0.3206
Si	-288.4517	-288.4513	0.4028
S	-396.9984	-396.9982	0.1758
Ar	-526.3012	-526.3012	0.0324
H_2	-1.1722	-1.1721	0.0605
C_2H_2	-76.8945	-76.8944	0.0663
C_2H_4	-78.1630	-78.1628	0.2654
CH_4	-40.2082	-40.2081	0.0883
CO	-112.7440	-112.7439	0.1267
CO_2	-187.7068	-187.7066	0.1756
CH_2O	-113.9527	-113.9525	0.2078
CHN	-92.9260	-92.9258	0.1871
HF	-100.0409	-100.0409	0.0531
N_2	-108.9666	-108.9665	0.1651
NH_3	-56.2977	-56.2976	0.1216
NaCl	-620.8285	-620.8285	0.0301
LiH	-7.9921	-7.9920	0.0256
Li_2	-14.8343	-14.8343	0.0292
Be_2	-29.0617	-29.0616	0.0482
LiF	-106.9356	-106.9356	0.0511
LiCl	-466.5985	-466.5985	0.0292
BH_3	-26.4238	-26.4237	0.0971
F_2	-198.6967	-198.6963	0.4158
O_3	-224.4376	-224.4373	0.2726
HNO	-129.8671	-129.8669	0.1852
HCF	-137.7739	-137.7738	0.1624
CHCl	-497.4301	-497.4299	0.1961
Avr.			0.1478

Table 4.1: Total energies for a selection of closed shell systems calculated using LDA and constrained LDA (cLDA), as well as the difference between these two energies. All calculations use a cc-pVTZ orbital basis set and the constrained calculations use an uncontracted cc-pVDZ auxiliary basis.

System	E^{NPOS} (Ha)	E^{POS} (Ha)	ΔE (Ha)	$\log(\Delta E)$
He	-2.8716	-2.8690	-2.58E-03	-2.59
LiH	-7.9917	-7.9915	-2.19E-04	-3.66
Be	-14.5201	-14.5201	-3.56E-06	-5.45
C	-37.5122	-37.5122	-1.22E-06	-5.91
O	-74.5765	-74.5765	-4.00E-07	-6.40
HF	-100.0337	-100.0337	-1.54E-07	-6.81
Ne	-128.4161	-128.4161	-2.00E-08	-7.70

Table 4.2: Ground state total energies obtained for cLDA both with (POS) and without (NPOS) the positivity constraint enforced on the effective potential. In the third and forth columns a measure of the energy difference $\Delta E = E^{\mathrm{NPOS}} - E^{\mathrm{POS}}$ is given. All calculations use a cc-pVTZ orbital basis set and the constrained calculations use an uncontracted cc-pVDZ auxiliary basis. For the positivity enabled calculations a penalty of $\Lambda = 100$ a.u. was used.

ilar and within $< 10^{-5}$ Ha. Energies obtained with the positivity constraint are also expectedly slightly higher due to the additional constraint placed on minimisation (see Table 4.2). It is clear from the results presented that the total energies calculated using the constrained method are indeed slightly higher in energy than those calculated from an unconstrained DFT calculation, as we would expect. We can also note that the average difference between the two energies is of the order of $0.1 - 0.2$ mHa, hence we can conclude that imposing these additional constraints on the system has very little impact on its total energy and the converged solution is the same as a standard DFT calculation.

4.2.4 Ionisation energies

In theory, using the Koopmans' theorem of DFT (see § 4.1.1), the ionisation energy is given exactly by minus the value of the HOMO eigenvalue, ϵ_{ho}. For approximate functionals however, only approximations for the ionisation energy can be obtained. Therefore, the accuracy of the KS HOMO eigenvalue to the experimental result for a systems ionisation energy is a good measure of functional accuracy.

Table 4.3 provides a comparison of the experimental ionisation energies listed in the NIST computational chemistry database [114] to the KS HOMO eigenvalue

obtained from a standard, unconstrained LDA calculation and the constrained LDA method. For the set of systems presented in this table, the average error between the experimental and calculated values from unconstrained LDA is 4.16 eV compared to an average error of 0.62 eV for constrained LDA. Therefore, it is clear that the self-interaction correction enforced through the $N-1$ screening charge constraint yields a dramatic improvement in calculated ionisation energies, without increasing considerably the total energy.

4.2.5 Electron affinities

After investigating ionisation energies, the next logical step is to investigate accuracy of electron affinities calculated using the constrained method. Unlike for ionisation energies, the electron affinity of the system is not given exactly by the KS LUMO eigenvalue (a relationship does exist and will be explored in § 6). Therefore, when comparing ϵ_{lu} against the affinity of a system, accurate results for (semi-)local density functional approximations (even with the constraint of $N-1$), are not to be expected.

There is however an alternative way to obtain more accurate affinity using the constrained method. As presented in Table 4.4, calculated ionisation energies from the constrained method are systematically closer to experimental results than the unconstrained results. It is therefore possible to invoke the Koopmans' theorem for DFT for the $N+1$ electron system, and approximate its ionisation energy as the electron affinity of the N electron system, where:

$$A \approx \epsilon_{\mathrm{ho}}^{N+1}. \tag{4.44}$$

Table 4.4 compares results for the electron affinities for a range of systems against the experimental values listed in the NIST computational chemistry database (CCCBDB) [114]. As with the ionisation energies, there is a clear improvement in the calculated values for electron affinities using the constrained method applied to the $N+1$ system, compared to those calculated using unconstrained LDA. For

System	I^{exp}	$-\epsilon_{\text{ho}}^{\text{LDA}}$	$-\epsilon_{\text{ho}}^{\text{cLDA}}$	ΔI^{LDA}	ΔI^{cLDA}
He	24.59	16.02	23.22	8.57	1.37
Be	9.32	6.04	8.96	3.28	0.36
C	11.26	6.27	9.85	4.99	1.41
O	13.62	9.21	13.80	4.41	-0.18
Ne	21.57	14.06	19.47	7.50	2.10
Mg	7.65	5.20	7.75	2.44	-0.10
Si	8.15	4.86	7.55	3.29	0.60
S	10.36	7.24	10.36	3.12	0.00
Ar	15.76	10.89	14.47	4.87	1.29
H_2	15.43	10.78	15.77	4.65	-0.34
C_2H_2	11.40	7.83	11.12	3.57	0.28
C_2H_4	10.51	7.83	10.47	2.68	0.04
CH_4	12.61	7.80	11.40	4.81	1.21
CO	14.01	9.60	13.17	4.41	0.84
CO_2	13.78	9.81	13.10	3.97	0.68
CH_2O	10.89	6.82	10.24	4.07	0.64
CHN	13.60	9.68	13.17	3.92	0.43
HF	16.03	10.30	14.82	5.73	1.22
N_2	15.58	10.89	14.49	4.69	1.10
NH_3	10.07	6.75	10.58	3.32	-0.51
LiH	7.90	4.79	7.53	3.11	0.37
Li_2	5.11	3.63	5.60	1.48	-0.49
LiF	11.30	6.82	9.88	4.48	1.42
LiCl	9.57	6.47	9.59	3.10	-0.02
BH_3	12.03	8.96	12.42	3.06	-0.40
F_2	15.70	10.14	13.37	5.56	2.32
O_3	12.53	8.69	11.18	3.84	1.35
HNO	10.10	6.15	9.58	3.95	0.52
HCF	10.06	6.27	9.58	3.79	0.48
Avr.				4.16	0.62

Table 4.3: Ionisation energies in eV for a selection of closed shell systems calculated using LDA and constrained LDA compared to experimental values obtained from the NIST computational chemistry database (CCCBDB) [114]. All of the calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVDZ auxiliary basis set.

System	A^{exp}	$-\epsilon^{\text{LDA}}_{\text{ho}[N+1]}$	$-\epsilon^{\text{cLDA}}_{\text{ho}[N+1]}$	ΔA^{LDA}	ΔA^{cLDA}
Li	0.618	-0.515	0.427	1.133	0.191
F	3.401	-0.892	2.304	4.293	1.097
Na	0.548	-0.491	0.534	1.039	0.014
Cl	3.613	0.332	2.731	3.281	0.882
OH	1.828	-1.543	1.096	3.371	0.732
CN	3.862	0.586	3.066	3.276	0.796
NH2	0.771	-1.947	0.410	2.658	0.301
HS	2.314	-0.398	1.765	2.712	0.549
ClO	2.276	-1.234	1.032	3.510	1.244
CNO	3.609	0.210	2.797	3.399	0.812
SCN	3.537	0.812	2.986	2.725	0.551
Avr.				2.942	0.652

Table 4.4: Electron affinities in eV for a selection of base open shell systems calculated using LDA and constrained LDA compared to experimental values obtained from the NIST computational chemistry database (CCCBDB) [114]. All of the calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

these calculations, an augmented orbital basis set was employed to more accurately represent the delocalised HOMO orbital. Results obtained from this updated version of the cLDA code without positivity compare favourably to previous results obtained for several of these ions [100, 112], providing more justification for the relaxation of the positivity constraint in practice.

One issue that is worth mentioning here is that the $N+1$ electron system of a closed-shell N electron system is always open-shell. This causes problems for the unpolarised DFAs used in both the unconstrained and constrained calculations discussed so far. Normally, one would resort to SDFT to calculate open-shell systems but clearly the current formulation is strictly for restricted DFT. This problem will be revisited in the next chapter, however at this stage only $N+1$ systems that are closed-shell are considered.

			Avr. time (s)			Ratio	
System	n_{orb}	n_{aux}	old	new	pos.	old/new	pos./new
He	15	7	7.04	0.97	101.60	7.26	104.74
Be	35	26	20.67	2.98	59.26	6.94	19.89
O	35	25	40.11	5.99	53.70	6.70	8.96
Ne	35	25	16.03	2.22	22.06	7.22	9.94
Si	39	39	42.54	7.02	61.06	6.06	8.70
Ar	39	38	18.10	3.18	24.92	5.70	7.84
H_2	30	14	26.37	5.21	146.32	5.06	28.08
LiH	50	33	53.88	12.78	1112.07	4.22	87.02
HF	50	32	47.40	10.94	73.84	4.33	6.75
O_2	70	50	256.75	70.01	409.86	3.67	5.85
CO	70	50	133.63	31.00	1107.14	4.31	35.71
F_2	70	50	127.74	35.80	221.86	3.57	6.20
NaCl	78	77	159.31	49.89	606.01	3.19	12.15
CH_2	65	39	284.46	96.04	773.68	2.96	8.06
HNO	85	57	339.26	128.15	671.24	2.65	5.24
CO_2	105	75	418.33	140.20	640.82	2.98	4.57
NH3	80	46	445.51	213.88	1329.60	2.08	6.22
C_2H_2	100	64	527.10	188.32	1050.54	2.80	5.58
CH_2O	100	64	694.42	309.77	1223.60	2.24	3.95

Table 4.5: Average full SCF cycle computation times for a range of small systems to highlight the effect of streamlining the old OEP routines and the relaxation of the positivity constraint. Calculations labelled as 'old' and 'new' were performed without the positivity constraint, whereas those labelled as 'pos.' were performed using the new streamlined routines and the positivity constraint enforced.

4.2.6 Computational optimisations

Since the original implementation of the constrained method, many improvements have been made to the HIPPO code to improve the computation speeds of constrained OEP calculations. Table 4.5 highlights the difference that these changes have made as a function of the total number of basis functions using calculations from a selection of systems. The changes made to the routines to optimise the code yielded between a 2-7 times reduction in overall SCF computation time for the example systems studied. Streamlining the routines resulted in a larger reduction in average full SCF time in smaller systems, where a larger percentage of time is spent outside of loops over basis sets relative to in much larger systems. As previously discussed, the relaxation of the positivity constraint also provides a

huge reduction in calculation times, particularly in systems with a small number of electrons where the $N-1$ constraint plays a more significant role. All of these benchmark calculations were ran serially on a MacBook Pro (2021 model, M1 Max chip, 32GB memory), and each value averaged over 5 runs.

4.3 Summary

In this chapter, the constrained minimisation method was discussed and improvements were presented. It is clear that constraining the screening charge of a system to $N-1$ can correctly reproduce the asymptotic behaviour as the exact KS potential. Using this method systematic improvements to both ionisation energies and electron affinities can be made, with minimal impact of total energies. Further optimisations were also made to the HIPPO code to increase computational efficiency.

The relaxation of the positivity constraint previously enforced in the constrained minimisation method was shown to yield excellent convergence in systems with more than a couple of electrons, and for modest auxiliary basis sets. By lifting this constraint, a major computational bottleneck can therefore be avoided without incurring any significant accuracy loss in calculated ground state total energies, xc potentials, and orbital energies.

Chapter 5

Kohn-Sham DFT equations for open-shells

So far, the main focus of this thesis has been solving for the electronic structure of closed-shell systems. In these systems, every electron is paired, meaning the exchange-correlation potential can be approximated equally well by the potential arising from the spin-up electrons or the spin-down electrons, using an unpolarised exchange-correlation energy functional.

Open-shell systems on the other hand pose an additional level of complexity because they contain unpaired electrons. Using the standard, unpolarised approximations for the exchange-correlation functional yield poor results for these systems, and normally spin-DFT is used to more accurately model these systems.

In principle, however, DFT is a theory that can be applied to any electronic system, so why is it then that commonly used functionals in DFT yield poor results for any system with unpaired electrons compared to spin-DFT? Furthermore, if there is no external magnetic field applied to an electronic system, both DFT and spin-DFT should yield the same ground state density and the same total energy; and therefore the same value of any observable. In this limit, both DFT and spin-DFT are formally exact theories, but thus far no definitive link between the two has been established.

This chapter begins by addressing this apparent disconnect, where I derive a new set of generalised Kohn-Sham DFT equations for open-shell systems. I will then show how this new set of equations can provide a direct link to the Kohn-Sham equations of spin-DFT; defining in the process the first exact functional for an observable in DFT other than the density and total energy.

After establishing this link, I will discuss the long-held belief in the field that spin-DFT approximations are intrinsically more accurate than DFT approximations, by introducing and correcting an error that contaminates unpolarised DFT approximations. This involved the development of a new class of 'implicit' density functionals, and these new functionals are implicit functionals of the electronic density, they also require an OEP equation to be solved.

To implement these functionals into the HIPPO code, I worked extensively with Timothy Callow to adjust the constrained OEP formalism from the previous chapter to allow the separation of the two spin-channels; the details of which will be discussed. This collaborative work has been published in Ref. [115].

Finally, results I obtained from these new functionals are examined and compared against both spin-DFT and regular unpolarised DFT functionals, along with a discussion of how these new functionals provide a systematic improvement over any unpolarised density functional approximation for open-shell systems without incurring any accuracy loss for closed-shell systems. I then present several approximations that can be made to bypass the OEP scheme but retain a high level of accuracy in total energies and potentials calculated.

5.1 Theory

5.1.1 The missing link between exact formulations of DFT and spin-DFT

As was just mentioned, in the absence of any external magnetic fields both DFT and spin-DFT should yield (for the exact functional) exactly the same electronic density and total energy for any electronic system. The equality of these two theories in this limit has not however been explicitly derived, and in general, spin-DFT is the preferred method of choice when dealing with any systems that contain unpaired electrons. This is because the approximate functionals employed in the formalisms are fundamentally different. In DFT, the most commonly used xc-functionals are modelled on the homogeneous spin-unpolarised electron gas, whereas in spin-DFT they are modelled on the spin-polarised electron gas. The two fundamental variables of spin-DFT, the spin-up and spin-down densities ($\rho^{\uparrow}$ and $\rho^{\downarrow}$), compared to the single variable in DFT, the total density, ρ, are thought to provide more flexibility in the construction of exchange-correlation energies. This has been summarised by Parr and Yang [116], where they write that for (semi-)local density functional approximations "the exchange-correlation energy of the electrons is approximated locally by results for the homogeneous spin-compensated electron gas. Such a procedure is not appropriate for systems with unpaired electrons, like open-shell molecules. A better description for such systems will be obtained through the use of the exchange-correlation energy of the homogeneous spin-polarised electron gas". It is the density functional approximations therefore that introduce the difference in the calculated values for observables obtained from the two theories, and this explains the origin of the belief that spin-DFT yields more accurate results for open-shell systems.

Before discussing the construction and implementation of DFT approximations that can yield (near) spin-DFT accuracy, first, the disconnect between DFT and

spin-DFT in the limit of zero applied magnetic field will be explored. Below, a new set of Kohn-Sham equations for DFT are derived that reduce to the Kohn-Sham equations of spin-DFT. This set of equations, therefore, provides a connection between the two theories in this limit.

5.1.2 The DFT-generalised Kohn-Sham scheme

This derivation begins with the universal energy functional from the constrained search formulation of Levy and Lieb [36–38] that was introduced earlier in this thesis:

$$F[\rho] = \min_{\Psi \to \rho} \langle \Psi | \, \hat{T} + \hat{V}_{\mathrm{ee}} \, | \Psi \rangle \, . \tag{5.1}$$

The state that minimises the universal functional, Ψ_ρ, depends on the total density ρ. This total density will be the ground state density of a fully-interacting electronic system bound by a local potential, v. Therefore, the minimising state, Ψ_ρ, is exactly the ground state wave-function of this system (assuming that we limit our search to only consider densities that are pure-state v-representable).

For collinear applied magnetic field, including the case of $B = 0$, the number of spin-up and spin-down electrons, $N^\uparrow$ and $N^\downarrow$, are good quantum numbers for the Hamiltonian (though strictly, this is only the case for an insulator), and it follows therefore that the minimising state Ψ_ρ can be chosen to have integer values for $N^\uparrow$ and $N^\downarrow$. In standard KS DFT, the KS reference state, Φ_ρ, is determined by minimising the non-interacting kinetic energy functional $T_{\mathrm{s}}[\rho]$, and it too can be chosen to have integer values of $N^\uparrow$ and $N^\downarrow$. This minimisation is similar to the minimisation of the universal functional and later the outcomes of these two will be compared.

Inspired by the approach of Levy [36], the total minimisation of the universal functional is split into two separate minimisation steps, which have the same overall minimum as the single minimisation:

$$F[\rho] = \min_{(\rho^\uparrow, \rho^\downarrow) \to \rho} \left[\min_{\Psi \to (\rho^\uparrow, \rho^\downarrow)} \langle \Psi | \, \hat{T} + \hat{V}_{\mathrm{ee}} \, | \Psi \rangle \right] . \tag{5.2}$$

The first minimisation inside the angled brackets is performed over all possible states Ψ that have a spin-density $(\rho^{\uparrow}, \rho^{\downarrow})$. This is then followed by the second minimisation that is performed over all spin-densities that sum to a total density ρ. It is important to note here that the minimising total density in this two-stage minimisation is exactly the same as the minimising total density of the second minimisation of equation 5.1, as both have the same overall minimum. This is an interesting result, as it confirms that not only is the minimising total density obtained from the two-step minimisation process exactly the true ground-state density of the real interacting system, but also that the minimising spin-densities are equal to the spin-densities of the exact interacting system; which are themselves functionals of the total density. Hence, here we obtain the elusive density functional of the spin-density in DFT. This is a very significant result, as here we have in fact defined the first exact functional within DFT for a quantity other than the density or the total energy.

Taking a closer look at the first minimisation step in 5.2, it is clear that it actually defines the universal energy functional for spin-DFT:

$$F[\rho] = \min_{(\rho^{\uparrow},\rho^{\downarrow})\to\rho} F[\rho^{\uparrow}, \rho^{\downarrow}], \tag{5.3}$$

where:

$$F[\rho^{\uparrow}, \rho^{\downarrow}] = \min_{\Psi\to(\rho^{\uparrow},\rho^{\downarrow})} \langle\Psi|\, \hat{T} + \hat{V}_{\mathrm{ee}}\, |\Psi\rangle\,. \tag{5.4}$$

It can therefore be concluded that the universal functional on DFT can be obtained by an extra minimisation of the universal functional of spin-DFT over all spin-densities with the same total density.

Expanding the spin-DFT universal functional in terms of its component functionals yields:

$$F[\rho] = \min_{(\rho^{\uparrow},\rho^{\downarrow})\to\rho} \Big[T_{\mathrm{s}}[\rho^{\uparrow}, \rho^{\downarrow}] + E_{\mathrm{xc}}[\rho^{\uparrow}, \rho^{\downarrow}]\Big] + U[\rho], \tag{5.5}$$

where $T_{\mathrm{s}}[\rho^{\uparrow}, \rho^{\downarrow}]$ and $E_{\mathrm{xc}}[\rho^{\uparrow}, \rho^{\downarrow}]$ are the known kinetic and exchange-correlation functionals of spin-DFT, respectively.

At the minimum, by definition, the functional derivative must vanish under the constraint that the total density, ρ, must integrate to $N^{\uparrow}+N^{\downarrow}$. With these constraints in hand, the following Euler-Lagrange equation can be derived that determines the minimising spin-DFT KS system with the minimising spin-density:

$$\left.\frac{\partial T_{\mathrm{s}}[\rho^{\uparrow},\rho^{\downarrow}]}{\partial\rho^{\sigma}(\mathbf{r})}\right|_{\rho^{\uparrow}_{\rho},\rho^{\downarrow}_{\rho}}+\left.\frac{\partial E_{\mathrm{xc}}[\rho^{\uparrow},\rho^{\downarrow}]}{\partial\rho^{\sigma}(\mathbf{r})}\right|_{\rho^{\uparrow}_{\rho},\rho^{\downarrow}_{\rho}}+v_{\mathrm{eff}}[\rho](\mathbf{r})=0, \tag{5.6}$$

where $v_{\mathrm{eff}}[\rho](\mathbf{r})$ is a local potential consisting of a continuous set of Lagrange multipliers that enforces the total density constraint on the system.

The first term in the above equation, the functional derivative of the non-interacting kinetic energy functional with respect the spin density at $(\rho^{\uparrow}_{\rho},\rho^{\downarrow}_{\rho})$ is equal to minus the spin-dependent KS potential (within a constant). The second term is simply the spin-dependent exchange-correlation potential. The effective potential can also be broken up into the sum of the Hartree potential and a remainder potential. Therefore, the unrestricted KS potential in DFT with the correct minimising spin-density, also in DFT, is given by:

$$v^{\sigma}_{\mathrm{s}}[\rho^{\uparrow}_{\rho},\rho^{\downarrow}_{\rho}](\mathbf{r})=v^{\sigma}_{\mathrm{xc}}[\rho^{\uparrow}_{\rho},\rho^{\downarrow}_{\rho}](\mathbf{r})+v_{\mathrm{H}}[\rho](\mathbf{r})+v[\rho](\mathbf{r}). \tag{5.7}$$

Notice that all of the terms in the definition of the KS potential are functionals of the total density, not the spin-densities. Furthermore, the spin-densities, ρ^{σ}_{ρ}, are also implicit functionals of ρ. Using the HK theorem of spin-DFT [66], the remaining local potential $v[\rho](\mathbf{r})$ can be identified as the external potential of the fully-interacting system.

The system that is determined by this new set of KS equations at $B=0$ is a new system that will be termed the generalised DFT KS system. Despite the absence of a non-local Fock-exchange like term, this new set of equations are still an example of a GKS scheme [61]. The potential present in this new DFT GKS scheme is always local but spin-dependent. The key result here is not simply the reduction of the spin-DFT KS equations to the generalised DFT KS equations. This new minimising state is the Slater determinant $\Phi_{\rho^{\uparrow}_{\rho},\rho^{\downarrow}_{\rho}}$ which yields not only the correct

total density of the interacting system but also the correct spin-density. This is an important new result: the observable spin-density is a functional of the total density.

Now that this link between the two exact theories in this limit has been established, attention turns to constructing density functional approximations in DFT that can yield near spin-DFT accuracy for open-shell systems and to explaining why spin-DFT approximations are perceived as intrinsically more accurate than their unpolarised counterparts.

5.1.3 Implicit functionals for open-shell systems

Local and semi-local density functional approximations that only depend explicitly on the total electronic density, ρ, generally assume that the spin-up and spin-down densities of a system are equal. For closed-shell systems, where every occupied orbital contains two paired electrons, this assumption is valid. This changes however for systems with unpaired electrons, i.e. open-shell, where clearly the two spin-densities are not ever equal. In the case of closed-shell systems, the following relation for the total electronic exchange energy holds:

$$E_{\mathrm{x}}[\rho] = E_{\mathrm{x}}\left[\frac{\rho}{2}, \frac{\rho}{2}\right], \tag{5.8}$$

where the functionals $E_{\mathrm{x}}[\rho]$ and $E_{\mathrm{x}}\left[\frac{\rho}{2}, \frac{\rho}{2}\right]$ are the exchange energy functionals of DFT and spin-DFT respectively. As we are considering closed-shell systems, each of the spin-densities, $\rho^{\uparrow}$ and $\rho^{\downarrow}$, are equal to half of the total density ρ; hence they do not appear explicitly in the above spin-DFT exchange functional.

For an open-shell system, this relationship is only approximate, however, as it wrongly introduces partial mixing between the spin-up and spin-down densities. By inserting an error term into this relationship that accounts for this mixing, Equation 5.8 can be made exact for all systems:

$$E_{\mathrm{x}}[\rho] = E_{\mathrm{x}}\left[\frac{\rho}{2}, \frac{\rho}{2}\right] + G_{\mathrm{x}}, \tag{5.9}$$

where this error, G_{x}, termed the 'ghost exchange' error, is defined as the difference between the spin-unpolarised and spin-polarised exchange energies:

$$G_{\mathrm{x}} = E_{\mathrm{x}}\left[\frac{\rho}{2}, \frac{\rho}{2}\right] - E_{\mathrm{x}}[\rho^{\uparrow}, \rho^{\downarrow}] \tag{5.10}$$

An important point to stress before continuing is that this error is not caused by representing the exchange energy in an approximate manner via a (semi-)local density functional approximation. Even the exact form of the exchange energy will be contaminated with this error within DFT for open-shells if one uses $\rho^{\uparrow} = \rho^{\downarrow} = \rho/2$. The origin of the ghost exchange is solely due to spin-mixing in the exchange energy incurred by the mistreatment of an open-shell system as effectively closed-shell.

To visualise the magnitude of the ghost exchange error, its definition can be linearly interpolated between the edge cases of completely spin-up and spin-down polarised systems:

$$G_{\mathrm{x}}(\omega) = E_{\mathrm{x}}[\rho_{\omega}, \rho_{\omega}] - E_{\mathrm{x}}[\rho^{\uparrow}, \rho^{\downarrow}], \tag{5.11}$$

where $\rho_{\omega} = (1-\omega)\rho^{\uparrow} + \omega\rho^{\downarrow}$, $0 \leq \omega \leq 1$. In the top subplot of Figure 5.1, the magnitude of the ghost exchange error for several atoms is plotted as a function of ω. It is clear that when there is no mixing between the spin-up and spin-down densities, i.e. at $\omega = 0$ and $\omega = 1$, the value of the ghost exchange error is zero, and that at $\omega = 0.5$, where the maximal amount of mixing is present, the G_{x} is also maximised. The bottom subplot in this figure highlights the relative contribution of the ghost exchange error to the total energy of the system. As the number of electrons increases, the percentage of the total energy contributed by the ghost exchange error decreases.

Having identified the ghost exchange error, the focus changes to the construction of density functional approximations that can correct for it. Starting from the Fock expression for the exact exchange energy in KS DFT:

$$E_{\mathrm{x}}[\rho^{\uparrow}, \rho^{\downarrow}] = -\frac{1}{2}\sum_{\sigma}^{\uparrow\downarrow} \iint \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{r}' \frac{|\rho^{\sigma}(\mathbf{r}, \mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|}, \tag{5.12}$$

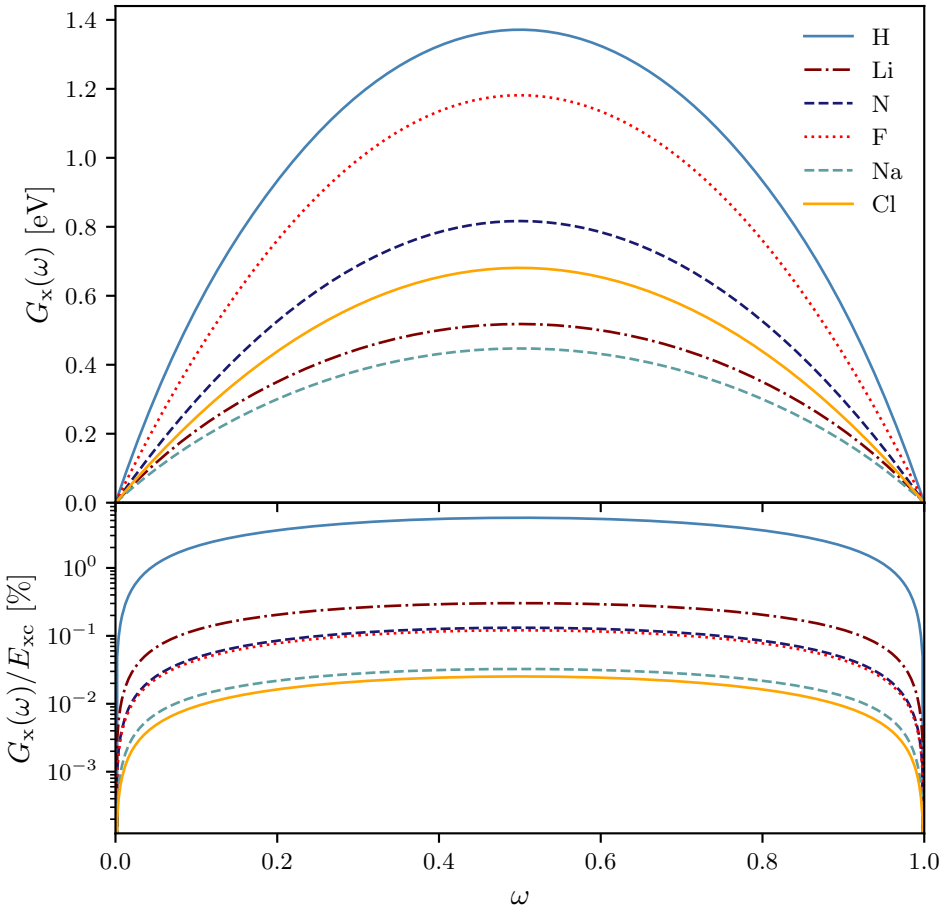


Figure 5.1: The magnitude of the ghost exchange error in LSDA and its relative percentage of the total exchange-correlation energy as a function of ω for a selection of atoms to get an idea of its relative contribution.

where $\rho^\sigma(\mathbf{r}, \mathbf{r}')$ is the one-body reduced density matrix for the spin-σ Kohn-Sham orbitals, it follows that the exact exchange functional can be separated into two terms, one related to each spin-channel:

$$E_x[\rho^\uparrow, \rho^\downarrow] = E_x[\rho^\uparrow, 0] + E_x[0, \rho^\downarrow] \tag{5.13}$$

Any density functional that depends only on the total electronic density and its respective derivatives will not satisfy Equation 5.13, and will therefore contain some amount of spin-mixing. If however the exchange energy is modelled in terms of the KS spin-density, the relation is always satisfied. At first glance, this approach may

not appear to be within DFT. However, as the KS spin-densities are treated as functionals of the total density, the overall functional is an implicit functional of the total density; therefore remaining within the scope of DFT and not spin-DFT. The exchange energy for these implicit density functionals (iDFs) can be expressed as:

$$E_{\mathrm{x}}[\rho] = E_{\mathrm{x}}^{\mathrm{iDF}}[\rho] = E_{\mathrm{x}}^{\mathrm{sDF}}[\rho^{\uparrow}[\rho], \rho^{\downarrow}[\rho]], \tag{5.14}$$

where $E_{\mathrm{x}}^{\mathrm{sDF}}$ is a spin-polarised density functional (sDF), where the spin-densities are themselves functionals of the total density (with the exact $E_{\mathrm{x}}^{\mathrm{sDF}}$ given by the Fock exchange-energy).

So far, only the exchange part of exchange-correlation functionals has been considered. Unlike with electron exchange, where the exact functional $E_{\mathrm{x}}[\rho]$ is equal to $E_{\mathrm{x}}^{\mathrm{iDF}}[\rho]$, the correlation energy functional does not satisfy a similar relation, as correlation between opposite spin electrons can occur. The exact correlation energy functional can therefore be written as:

$$E_{\mathrm{c}}[\rho] = E_{\mathrm{c}}^{\mathrm{iDF}}[\rho] + \Delta E_{\mathrm{c}}[\rho], \tag{5.15}$$

where $\Delta E_{\mathrm{c}}[\rho]$ is a correlation remainder functional and the implicit correlation energy functional is given by:

$$E_{\mathrm{c}}^{\mathrm{iDF}}[\rho] = E_{\mathrm{c}}^{\mathrm{sDF}}[\rho^{\uparrow}[\rho], \rho^{\downarrow}[\rho]]. \tag{5.16}$$

Up until now everything has been exact. However, as in general the remainder term $\Delta E_{\mathrm{c}}[\rho] \neq 0$, the following treatment becomes approximate if it is neglected. Luckily, spin-polarised correlation functionals tend to provide only a small improvement in accuracy over unpolarised ones, unlike for exchange where the ghost exchange error is far more significant.

Combining expressions 5.14 and 5.16 for the exchange and correlation functionals respectively, an accurate approximation for the xc functional is:

$$E_{\mathrm{xc}}[\rho] \approx E_{\mathrm{xc}}^{\mathrm{iDF}}[\rho] = E_{\mathrm{xc}}^{\mathrm{sDF}}[\rho^{\uparrow}[\rho], \rho^{\downarrow}[\rho]]. \tag{5.17}$$

It follows therefore that the total electronic energy functional containing the new implicit exchange-correlation functional is:

$$E^{\mathrm{iDF}}[\rho] = T_{\mathrm{s}}[\rho] + \int \mathrm{d}\mathbf{r}\; v_{\mathrm{en}}(\mathbf{r})\rho(\mathbf{r}) + E_{\mathrm{H}}[\rho] + E_{\mathrm{xc}}^{\mathrm{iDF}}[\rho], \tag{5.18}$$

where $T_{\mathrm{s}}[\rho]$, $v_{\mathrm{en}}(\mathbf{r})$, and $E_{\mathrm{H}}[\rho]$ are the familiar Kinetic energy functional, external potential, and Hartree energy functional, respectively.

5.1.4 The OEP equation for v_{xc} as an implicit density functional for open-shells

As for all implicit exchange-correlation functionals, the xc-potential, $v_{\mathrm{xc}}[\rho](\mathbf{r})$, must be calculated via an OEP equation. For this new kind of implicit functionals, the OEP equation takes the following form:

$$\int \mathrm{d}\mathbf{r}'\chi(\mathbf{r},\mathbf{r}')v_{\mathrm{xc}}[\rho](\mathbf{r}') = \sum_{\sigma}^{\uparrow\downarrow}\int \mathrm{d}\mathbf{r}'\chi^{\sigma}(\mathbf{r},\mathbf{r}')v_{\mathrm{xc}}^{\sigma}[\rho^{\uparrow}[\rho],\rho^{\downarrow}[\rho]](\mathbf{r}'), \tag{5.19}$$

where:

$$v_{\mathrm{xc}}^{\sigma}[\rho^{\uparrow}[\rho],\rho^{\downarrow}[\rho]](\mathbf{r}') = \left.\frac{\partial E_{\mathrm{xc}}^{\mathrm{sDF}}[\rho^{\uparrow},\rho^{\downarrow}]}{\partial\rho^{\sigma}(\mathbf{r}')}\right|_{\rho^{\uparrow}=\rho^{\uparrow}[\rho],\,\rho^{\downarrow}=\rho^{\downarrow}[\rho]}, \tag{5.20}$$

and $\chi(\mathbf{r},\mathbf{r}') = \sum_{\sigma}^{\uparrow\downarrow}\chi^{\sigma}(\mathbf{r},\mathbf{r}')$, where the spin-$\sigma$ response function is given by:

$$\chi^{\sigma}(\mathbf{r},\mathbf{r}') = 2\sum_{i=1}^{N^{\sigma}}\sum_{a=N^{\sigma}+1}^{\infty}\frac{\phi_i(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a(\mathbf{r}')}{\epsilon_i-\epsilon_a}. \tag{5.21}$$

It is important to note how this above formulation reduces to the same form of the OEP equation from § 4 in the case of a closed-shell system. In this case, the total response function is simply given by either $2\chi^{\uparrow}$ or $2\chi^{\downarrow}$, and the two spin potentials, $v_{\mathrm{xc}}^{\uparrow}(\mathbf{r})$ and $v_{\mathrm{xc}}^{\downarrow}(\mathbf{r})$ are equal. Therefore, we expect this new OEP formulation to yield identical results for closed-shell systems as the previously discussed implementation of the constrained OEP method.

Furthermore, if a system is completely spin-polarised, i.e. $\rho^{\sigma} = \rho$, $\rho^{\bar{\sigma}} = 0$ for $\sigma,\bar{\sigma} \in \{\uparrow,\downarrow\}$, the spin-$\bar{\sigma}$ response function is zero. Therefore, the solution of the OEP equation will simply be the spin-σ xc-potential.

For truly open-shell systems, it is clear that the two spin-dependent response functions will not be equal, and that they will act as weighting factors on each of the two spin-channels. Several approximations for this weighting that allow one to bypass the solution of the OEP equation will be explored later in this chapter.

This new method for constructing the ghost exchange corrected implicit density functionals can be applied to any existing (semi-)local approximation. For this work, both the iLDA and iPBE functionals were implemented into the HIPPO code, by extending the constrained OEP code. This involved changing the construction of the right-hand side of the OEP equation, $b(\mathbf{r})$, as the inversion procedure and left-hand side of the equation remain unchanged from the improved constrained OEP routine discussed in the previous chapter.

The implementation of the iPBE functional presented in [117] has been modified to prevent numerical instabilities that occurred when expanding the spin-xc potentials in certain auxiliary basis sets.

5.2 Results

5.2.1 Convergence of the implicit OEP potential

It is important to first establish whether solving the implicit OEP equation yields well-behaved potentials. We can begin by studying the convergence of the implicit OEP with the addition of a complement term. It is worth noting that results presented in this chapter, and in those that follow, were obtained without enforcing the positivity constraint discussed in § 4 (unless otherwise stated). Figure 5.2 shows the calculated xc potentials for NH_2 and CN obtained using a fixed augmented cc-pVTZ orbital basis set, with an auxiliary basis set of increasing size, from uncontracted cc-pVDZ to QZ from top to bottom. The expected behaviour of the potentials with decreasing values of α is evident; with well-behaved potentials appearing from $\alpha = 10^{-2}$. Some minor deviations around the atoms are observed

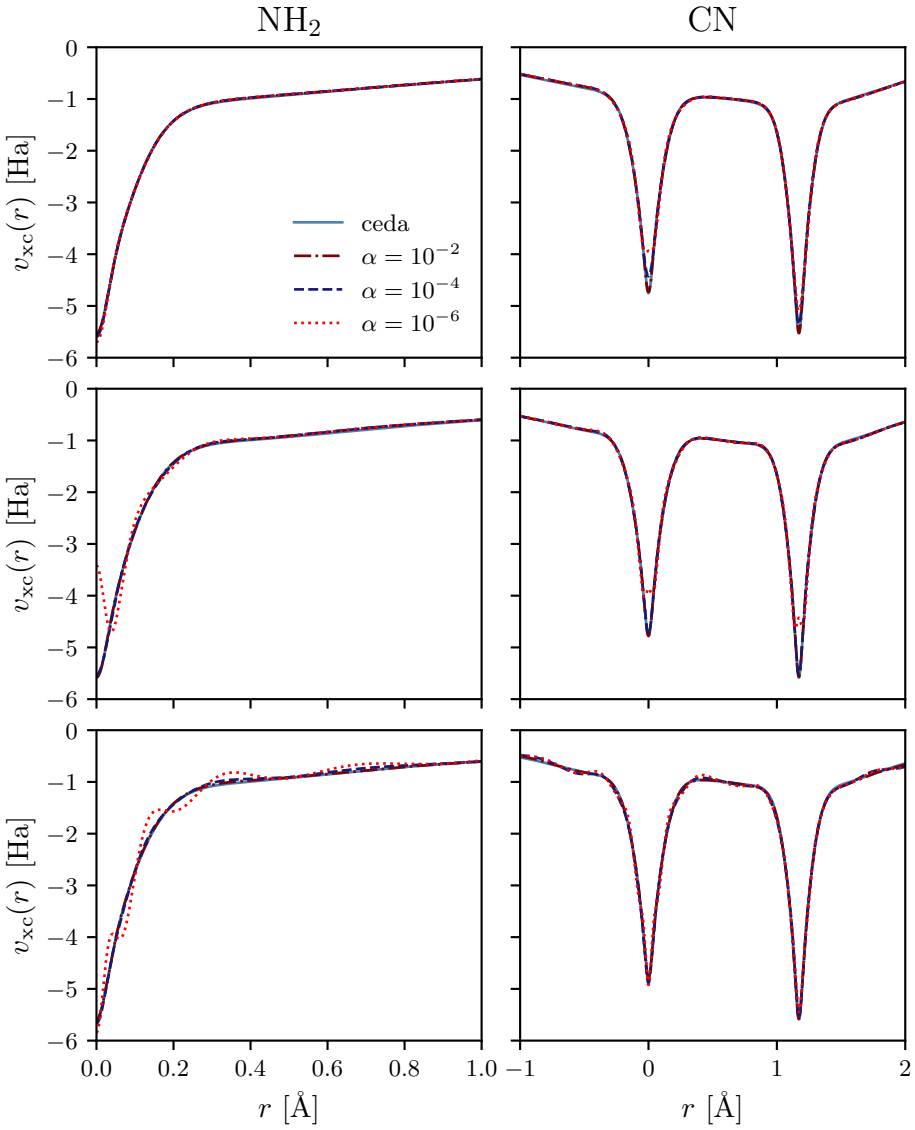


Figure 5.2: Convergence of the implicit OEP xc potential for NH_2 and CN for a range of values of α. The orbital basis set was fixed as augmented cc-pVTZ for all calculations, with an uncontracted (top) cc-pVDZ, (middle) cc-pVTZ, and (bottom) cc-pVQZ auxiliary basis set.

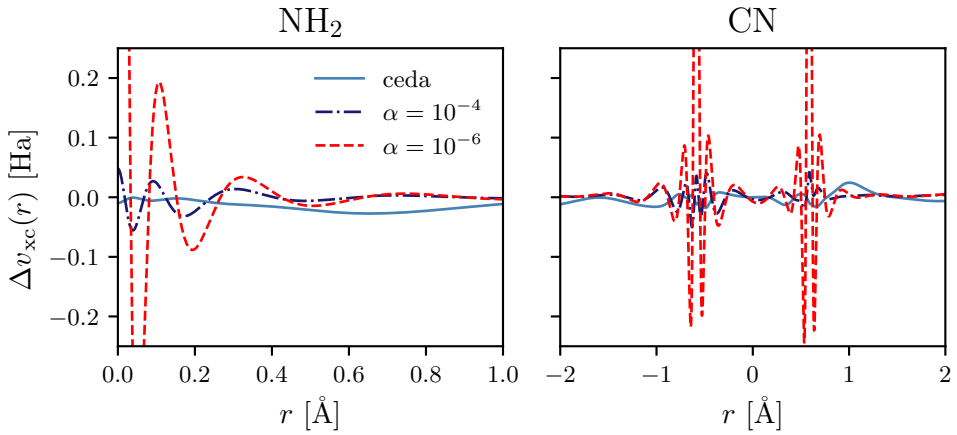


Figure 5.3: Differences between the iLDA potentials for NH_2 and CN calculated using CEDA, $\alpha = 10^{-4}$ and $\alpha = 10^{-6}$ and the potentials obtained with $\alpha = 10^{-2}$. These calculations were all performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

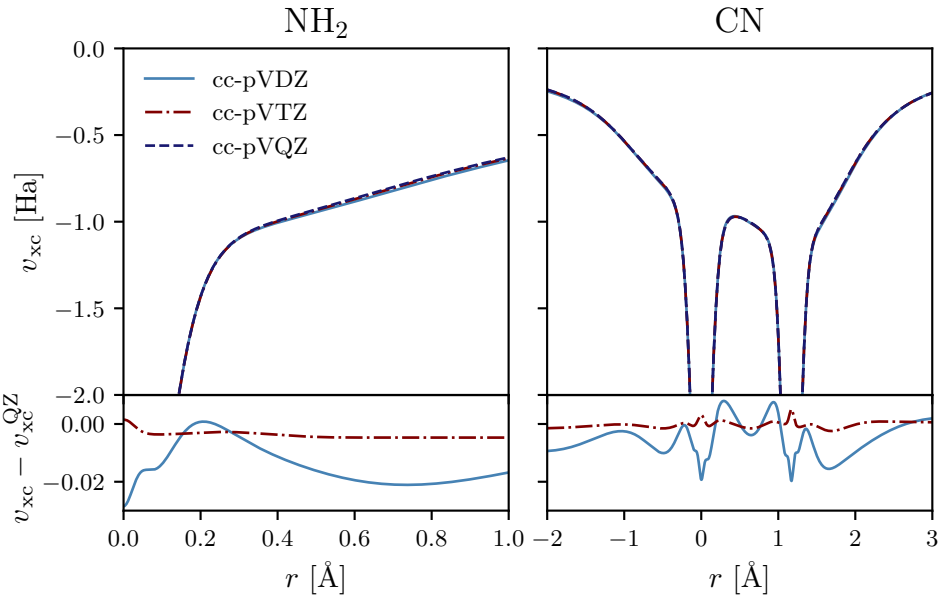


Figure 5.4: Convergence of the implicit LDA xc potential for NH_2 with varying orbital basis and fixed uncontracted cc-pVQZ auxiliary basis set. The bottom subplots show the difference between cc-pVDZ and TZ potentials relative to the QZ result. All calculations use a value of $\alpha = 10^{-2}$.

at $\alpha = 10^{-4}$, and more pronounced oscillations at $\alpha = 10^{-6}$ in the uncontracted cc-pVQZ results. A value of $\alpha = 10^{-2}$, therefore, appears to be a good default choice for these systems; matching the observations made back in § 4.

It is also clear from Figures 5.2 and 5.3 that the CEDA approximation yields potentials that are very close (within ~ 0.05 Ha for each of the basis set combinations) to those obtained from the full OEP solution that includes a small complement term. This highlights an important point, in that the CEDA approximation can nearly always produce results that lie close to the full OEP solution. In addition, CEDA does not suffer from any of the mathematical issues of the OEP equation in finite basis implementations that can manifest with the addition of too small of a complement; these start developing in the bottom two plots of 5.2. Another benefit of the CEDA approximation is that it is also computationally less demanding than constructing the full right-hand side of the OEP equation via the complement method. Although, for the small systems studied in this thesis, the boost in efficiency is not significant, especially when compared to the relaxation of the positivity constraint of the constrained method.

As a final check, the convergence of the potentials with the orbital basis set should also be analysed. For the same two systems, Figure 5.4 plots the xc potentials obtained for a fixed uncontracted cc-pVDZ auxiliary basis set and a varying orbital basis set. For both systems, we observe very good convergence even with the cc-pVDZ basis, as the relative differences relative to the cc-pVQZ potential are small. Increasing the orbital basis to cc-pVTZ yields excellent convergence with only minor variations to the potential occurring with further increases in the size of the basis set.

5.2.2 Ground state energies

One of the key comparisons that can be made between these new implicit density functional approximations and their unpolarised counterparts is in the calculated

System	E^{LSDA}	E^{cLDA}	E^{iLDA}	ΔE^{cLDA} (mHa)	ΔE^{iLDA} (mHa)
He	-2.8717	-2.8716	-2.8716	0.1209	0.1209
Be	-14.5202	-14.5201	-14.5201	0.0273	0.0273
C	-37.5140	-37.5136	-37.5136	0.3611	0.3611
NH	-54.8362	-54.8360	-54.8360	0.1930	0.1930
NH_3	-56.2977	-56.2976	-56.2976	0.1216	0.1216
O	-74.5801	-74.5799	-74.5799	0.1979	0.1979
C_2H_2	-76.8945	-76.8944	-76.8944	0.0663	0.0663
LiF	-106.9356	-106.9356	-106.9356	0.0511	0.0511
Ne	-128.4218	-128.4218	-128.4218	0.0297	0.0297
Mg	-199.3689	-199.3686	-199.3686	0.3206	0.3206
Si	-288.4517	-288.4513	-288.4513	0.4028	0.4028
S	-396.9984	-396.9982	-396.9982	0.1758	0.1758
Avr.				0.1723	0.1723

Table 5.1: Total ground state energies in Ha calculated using LSDA, cLDA, and implicit LDA for a few closed-shell systems. Energy differences are taken with respect to the LSDA result. All of the calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVDZ auxiliary basis set.

System	E^{LSDA}	E^{cLDA}	E^{iLDA}	ΔE^{cLDA} (mHa)	ΔE^{iLDA} (mHa)
H	-0.4962	-0.4622	-0.4962	34.0539	0.0545
Li	-7.3982	-7.3888	-7.3981	9.4288	0.0339
B	-24.4477	-24.4335	-24.4470	14.2164	0.7198
N	-54.1537	-54.1316	-54.1521	22.1303	1.5918
NH_2	-55.5935	-55.5739	-55.5919	19.5853	1.5913
OH	-75.3767	-75.3528	-75.3749	23.8587	1.7376
CN	-92.2029	-92.1881	-92.2014	14.7918	1.5140
F	-99.2844	-99.2559	-99.2827	28.4983	1.7364
Na	-161.6572	-161.6492	-161.6572	8.0358	0.0845
Al	-241.5719	-241.5636	-241.5714	8.3826	0.5705
P	-340.2326	-340.220607	-340.2316	12.0157	0.9749
Cl	-459.0034	-458.9886	-459.0023	14.7843	1.0580
Avr.				17.4818	0.9723

Table 5.2: Total ground state energies in Ha calculated using LSDA, cLDA, and implicit LDA for a selection of open-shell systems. Energy differences are taken with respect to the LSDA result. All of the calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVDZ auxiliary basis set.

total energy, and there are two separate classes of systems that should be compared.

First, the calculated total energies for a set of closed-shell systems can be compared. As the implicit OEP equation reduces to the regular closed-shell constrained OEP equation for systems with an equal number of spin-up and spin-down electrons, it is expected that total energies for these systems should be identical. In Table 5.1, a quick sanity check is performed for a few closed-shell systems to confirm that this is indeed the case; with calculated cLDA and iLDA energies being equal (to within any numerical noise). The average differences between these total energies and the calculated LSDA energies are also small (approximately 0.17mHa), with these slight derivations attributed to the additional constraint of $N-1$ on the screening density that has been enforced on the OEP solutions.

Having established that the implicit OEP method can reproduce the results of the constrained method for closed-shell systems, attention turns to the analysis of open-shell total energies. In Table 5.2, the ground state total energies obtained using the constrained and implicit methods are again compared to the LSDA energies; this time for a set of open-shell atoms/radicals. This table highlights a significant increase in the accuracy of calculated total energies for the iLDA functional compared to cLDA. The difference between the iLDA and cLDA energies is a measure of the ghost exchange error present in cLDA for these systems. By correcting for this error iLDA is able to yield total energies that are close to LSDA energies, with an average difference again reducing to the order of a millihartree. As before, this small energy difference is expected due to the additional constraints placed on the constrained minimisation in the implicit OEP implementation. An additional small error is also expected due to the treatment of the correlation energy functional in the same manner as exchange (see § 5.1.3).

A similar analysis can be performed on a small set of closed- and open-shell systems for the PBE functional. In Table 5.3, the ground state total energies determined using the spin-polarised PBE functional (SPBE) and the spin-unpolarised constrained (cPBE) and implicit PBE (iPBE) are presented. For the closed-shell sys-

System	E^{SPBE}	E^{cPBE}	E^{iPBE}	ΔE^{cPBE} (mHa)	ΔE^{iPBE} (mHa)
H	-0.4998	-0.4588	-0.4996	41.0534	0.3591
He	-2.8924	-2.8918	-2.8918	0.6441	0.6441
Li	-7.4619	-7.4511	-7.4606	10.7883	1.2351
NH	-55.0876	-55.0855	-55.0855	2.0180	2.0180
NH_2	-55.8320	-55.8050	-55.8271	26.9200	4.8233
NH_3	-56.5124	-56.5100	-56.5100	2.3578	2.3578
OH	-75.6833	-75.6512	-75.6783	32.0975	4.9605
Ne	-128.8527	-128.8507	-128.8507	1.9450	1.9450
Na	-162.1666	-162.1584	-162.1652	8.1271	1.3685

Table 5.3: Total ground state energies in Ha calculated using SPBE, cPBE, and iPBE for a selection of small systems. Energy differences are taken with respect to the SPBE result. All of the calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVDZ auxiliary basis set.

tems, again the constrained and implicit energies are identical as expected. For the open-shell systems, the iPBE functional shows a similar improvement compared to the unpolarised cPBE, however, the magnitude of the energy differences relative to the spin-polarised functional are slightly larger than the LDA results previously discussed.

5.2.3 Non-OEP approximations

Although the constrained and implicit formulations of the OEP equation allow for robust solutions to be obtained, they still invoke solving for the full-OEP solution which is always more computationally demanding than a regular DFT calculation, especially for very large auxiliary basis sets. It would be useful therefore to be able to develop accurate non-OEP approximations that correct for the ghost exchange error and provide near spin-DFT accuracy whilst avoiding this additional computational complexity. Furthermore, any non-OEP approximation could be implemented into an existing electronic structure code that does not include an OEP implementation with relative ease.

By inspecting the form of the spin-response functions given in Equation 5.21, we note that its denominator contains the difference between occupied and unoccupied

eigenvalues, $\epsilon_i - \epsilon_a$. Therefore, the weighting of each spin-response function on the overall OEP potential depends roughly on the inverse of the HOMO-LUMO gap, $\Delta^\sigma = \epsilon_{N^\sigma+1} - \epsilon_{N^\sigma}$, of the spin KS orbitals.

The observation provides a starting point for developing approximations that can bypass the OEP entirely. By introducing an open-shell system where $N^\uparrow > N^\downarrow$ (for consistency, the majority spin-channel with the greater number of electrons will be referred to as the 'up'-channel), there are three clear approximations that could be made:

1. $v_{\mathrm{xc}}(\mathbf{r}) \approx v_{\mathrm{xc}}^{\uparrow}(\mathbf{r})$
2. $v_{\mathrm{xc}}(\mathbf{r}) \approx v_{\mathrm{xc}}^{\downarrow}(\mathbf{r})$
3. $v_{\mathrm{xc}}(\mathbf{r}) \approx v_{\mathrm{xc}}^{w}(\mathbf{r}) = \left[\Delta^{\downarrow} v_{\mathrm{xc}}^{\uparrow}(\mathbf{r}) + \Delta^{\uparrow} v_{\mathrm{xc}}^{\downarrow}(\mathbf{r})\right] / \left(\Delta^{\uparrow} + \Delta^{\downarrow}\right)$

In the first approximation, the xc-potential is given by just the up-spin channels xc-potential. As $\Delta^\uparrow < \Delta^\downarrow$, the up-spin channel spin-response function will always dominate in the right-hand side of the OEP equation, making this an easily justifiable approximation. In Figures 5.5 and 5.6, it is clear that this approximation produces xc-potentials that are very close to the full-OEP solution. These figures also show that using this approximation yields potentials that are a distinct improvement over the spin-unpolarised, or ghost exchange contaminated, solutions. In addition, total energies obtained in this manner, as seen in Table 5.4 are within close agreement with the full-OEP energies; especially when compared to those calculated from spin-unpolarised LDA.

The second approximation assumes that the xc-potential is instead given by just the spin xc-potential arising from the down-spin (minority spin) channel. It is clear in Figures 5.5 and 5.6 that the potentials obtained for this approximation are far more inaccurate compared to using just the up-channels potential. Finally, relative to the unpolarised potentials, this approximation yields potentials that are considerably worse. This can be easily explained by again considering the relative magnitude of

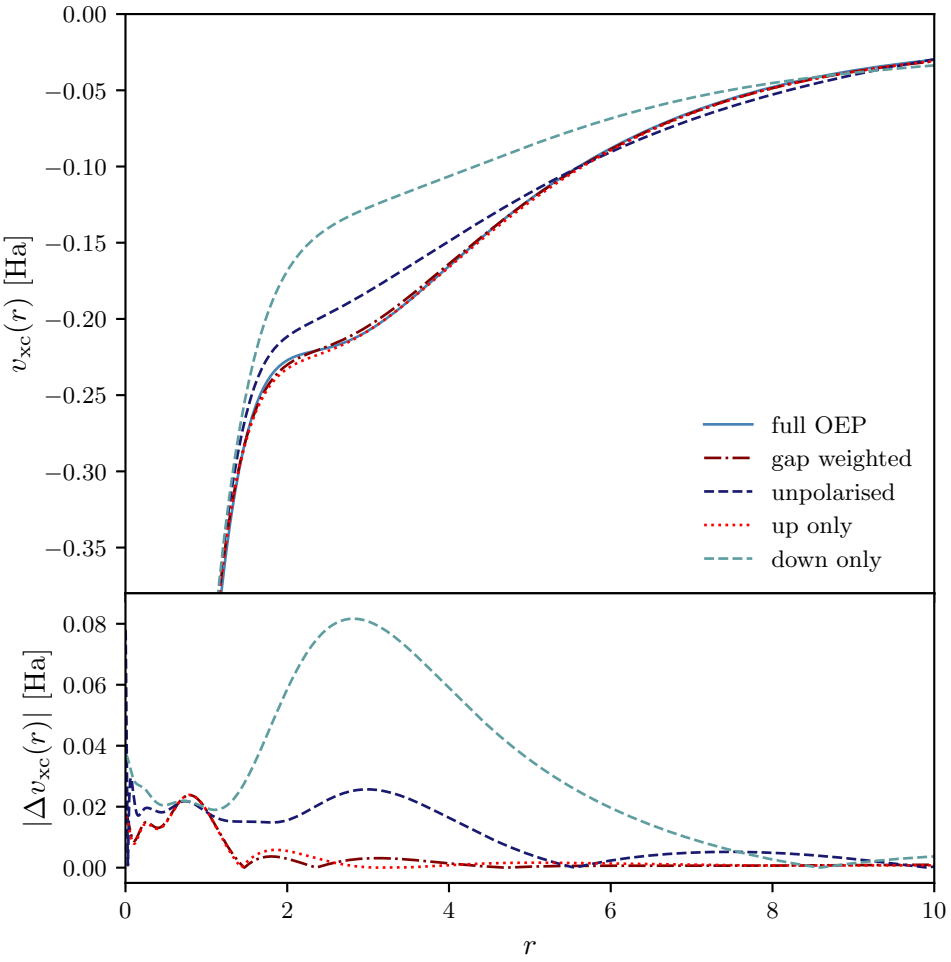


Figure 5.5: A comparison of the LDA potentials for Lithium obtained using non-OEP approximations to the implicit OEP equation compared to the full OEP solution with $\alpha = 10^{-2}$ and the unpolarised LDA solution. The bottom subplot shows the differences in the various potentials relative to the full-OEP solution.

the HOMO-LUMO gaps of the two spin-channels. As $\Delta^{\uparrow} < \Delta^{\downarrow}$, the down-channels contribution to the right-hand side of the OEP equation is far smaller than that of the up-channel and therefore provides a much smaller weighting to the overall xc-potential. Upon inspecting Table 5.4 again, it is clear that total energies are also less accurate than the previous approximation. However, it is interesting to note the relative errors in the total energies compared to spin-unpolarised DFT.

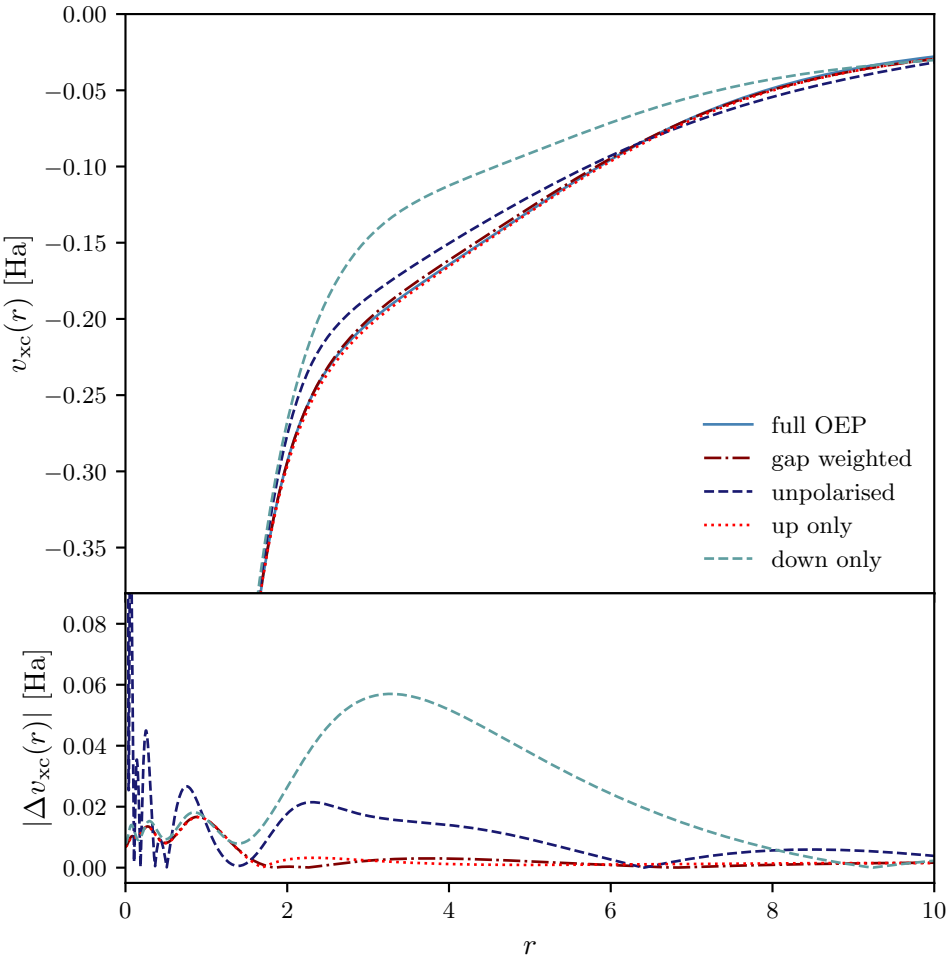


Figure 5.6: A comparison of the LDA potentials for Sodium obtained using non-OEP approximations to the implicit OEP equation compared to the full OEP solution with $\alpha = 10^{-2}$ and the unpolarised LDA solution. The bottom subplot shows the differences in the various potentials relative to the full-OEP solution.

This may seem surprising given the seemingly poor potentials calculated via this approximation, though by considering the definition of the ghost exchange error, which arises from the mixing between the two spin-channels, this approximation is ghost exchange free and are therefore 'relatively' accurate for total energies.

The third approximation takes a slightly more sophisticated form and attempts to directly include the observed inverse dependence of the spin-response functions

System	Method	E (Ha)	ϵ_{ho} (eV)	ΔE (mHa)	$\Delta\epsilon_{\text{ho}}$ (eV)
Li	LSDA	-7.3982	-3.5833	—	—
	iLDA	-7.3981	-5.8547	0.0339	—
	iLDA (v_{xc}^{w})	-7.3981	-3.5354	0.0345	0.0479
	iLDA ($v_{\text{xc}}^{\uparrow}$)	-7.3981	-3.5838	0.0348	-0.0005
	iLDA ($v_{\text{xc}}^{\downarrow}$)	-7.3964	-2.4103	1.7935	1.1730
	LDA	-7.3888	-3.2612	9.4288	0.3221
Na	LSDA	-161.6572	-3.4930	—	—
	iLDA	-161.6572	-5.7884	0.0845	—
	iLDA (v_{xc}^{w})	-161.6572	-3.4207	0.0882	0.0723
	iLDA ($v_{\text{xc}}^{\uparrow}$)	-161.6572	-3.5020	0.0866	-0.0090
	iLDA ($v_{\text{xc}}^{\downarrow}$)	-161.6565	-2.5124	0.7656	0.9806
	LDA	-161.6492	-3.1963	8.0358	0.2967
OH	LSDA	-75.3767	-8.7239	—	—
	iLDA	-75.3749	-12.7460	1.7376	—
	iLDA (v_{xc}^{w})	-75.3747	-8.5943	1.9367	0.1296
	iLDA ($v_{\text{xc}}^{\uparrow}$)	-75.3745	-8.3139	2.1431	0.4100
	iLDA ($v_{\text{xc}}^{\downarrow}$)	-75.3733	-8.3065	3.3562	0.4174
	LDA	-75.3528	-8.4275	23.8587	0.2964
NH_2	LSDA	-55.5935	-7.9001	—	—
	iLDA	-55.5919	-10.9760	1.5913	—
	iLDA (v_{xc}^{w})	-55.5911	-7.5210	2.4193	0.3791
	iLDA ($v_{\text{xc}}^{\uparrow}$)	-55.5915	-6.5714	2.0325	1.3287
	iLDA ($v_{\text{xc}}^{\downarrow}$)	-55.5903	-6.3158	3.2031	1.5843
	LDA	-55.5739	-7.0391	19.5853	0.8610

Table 5.4: Calculated ground state total energies and HOMO eigenvalues for four open-shell systems, obtained using LSDA, full and approximate iLDA and spin-unpolarised LDA, with differences relative to the LSDA results. The full iLDA results are OEP solutions that incorporate the screening charge constraint of $N-1$, whereas the approximate iLDA, LSDA and LDA solutions do not. All of the calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVDZ auxiliary basis set.

on their respective HOMO-LUMO gaps. By weighting the two potentials in this manner, one would expect to find potentials that are the most accurate of the three approximations; which is indeed observed in Figures 5.5 and 5.6 (although the improvement is minor over using just the up-channel due its dominance over the down-channel). Once again, the potentials obtained using this approximation are superior to the spin-unpolarised potentials. Total energies are comparable to the approximating the xc-potential as the spin-up potential, which again is to be

expected given the relative weighting of the two response functions.

From these observations, both the spin-up-only and gap-weighted approximations appear to provide both accurate potentials and energetics and offer a practical way of performing ghost exchange corrected calculations with no added cost or complexity compared to a regular spin-unpolarised DFT calculation.

An important point to note however is that employing a non-OEP approximation does not correct the asymptotic decay of the xc-potential enforced by the $N-1$ screening charge constraint in the full OEP calculation. Therefore, HOMO eigenvalues are expected to be in close agreement to those calculated using the unconstrained LDA functional; this can be observed in Table 5.4, where all of the approximations yield values of ϵ_{ho} much closer to the non-OEP LSDA results than to iLDA. Comparing the HOMO eigenvalues for the different methods and approximations to experimental ionisation energies (taken from the NIST CCCBDB [114]), the average differences obtained from LSDA, full-OEP iLDA and LDA are 2.66 eV, -0.26 eV and 3.10 eV, respectively. As the approximations discussed are unconstrained, it is clear therefore if one is interested in obtaining accurate orbitals eigenvalues for ionisation energies, full-OEP iLDA should be used.

5.3 Summary

In this chapter, a previously unknown link between spin-DFT and DFT was established by deriving a new set of generalised Kohn-Sham DFT equations in the limit of zero external magnetic field. The ghost exchange error was shown to be a major contributor to the misconception that spin-DFT approximations are intrinsically more accurate than DFT approximations, with minor contributions due to correlation neglected in this current work.

Through a collaborative effort, the implicit OEP method was implemented into the HIPPO code. This method provided a way to construct density functional

approximations that can yield spin-DFT accuracy for open-shell systems whilst firmly remaining within the realms of DFT.

Approximations to bypass the OEP framework were also introduced, providing an efficient way to calculate the electronic structure of open-shell systems that corrects a significant portion of the ghost exchange error.

Chapter 6

Ensemble OEP and the derivative discontinuity

Having now established a consistent OEP scheme that can accurately describe both closed-shell and open-shell systems, we can start to develop extensions to this theory that unlock the ability to calculate properties of the exact KS potential that are generally absent from most DFAs.

One of these properties is the discontinuous shift in the xc potential when the number of electrons in the system infinitesimally increases above an integer number, known commonly as the xc derivative discontinuity. Commonly used (semi-)local functionals, i.e. LDA and GGAs, completely lack a description of this derivative discontinuity.

In this chapter, we begin by presenting a brief overview of ensemble DFT (eDFT) and introduce in detail the derivative discontinuity and how it can be calculated. Next, we derive an OEP equation to determine the local potential that minimises the total ensemble energy for a system with a fractional average number of electrons. As in previous OEP schemes discussed, we will also discuss how the new ensemble OEP can be implemented into a finite basis set code.

Once we can solve this OEP equation for any fixed value of the ensemble weights, we then explore the properties of the calculated ensemble energies and potentials.

Finally, results obtained for the xc derivative discontinuity will be presented; calculated both directly from the gradient of the ensemble energies with varying electron number, and from the constant shift in the ensemble OEP potential observed within the vicinity of the system when an infinitesimal amount of an electron is added and removed. It is argued that the discontinuous change in the screening charge of the ensemble system as the average number of electrons surpasses an integer number is responsible for a large fraction of the overall xc derivative discontinuity.

6.1 Theory

6.1.1 Ensemble density functional theory

Kohn-Sham DFT was originally formulated as a theory for describing the ground state of an electronic system with a fixed integer N number of electrons [5, 6]. In its original formulation, it cannot however be used to accurately describe excited states. When we think about electronic excitations, it is important to note that there are two main types: neutral excitations (where the number of electrons is fixed) and charged excitations (where the number of electrons varies).

For neutral excitations, the most popular approach is to employ time-dependent-DFT (TD-DFT) [118, 119], whose main success is largely due to the relatively accurate results obtained using the adiabatic approximation. We will however focus on the case of charged excitations for the purposes of this brief review.

Charged excitations were first described by the extension of DFT to systems with a fractional electron number introduced by Perdew et al. [104]. However, implementation of this approach for the most commonly used (semi-)local functionals, i.e. LDA and GGAs, unfortunately, yields fundamental gaps that are far too small [120, 121]. This is largely due to the lack of the discontinuity in the derivative of the exact xc energy functional when the number of electrons in the system surpasses an integer number (see the next section for more detail) that is absent at this level of

approximation. As a result, to obtain more accurate fundamental gaps from DFT calculations, the most common approach is to invoke hybrid functionals [122], or to use a post-SCF technique such as *GW* [123]. Both of these approaches are however far more computationally demanding than (semi-)local DFAs (especially *GW*).

Ensemble density functional theory (eDFT) builds on the ideas first proposed in the extension of DFT to fractional electron number (which itself can be thought of as a type of ensemble theory). There have been several formulations of eDFT developed since its inception, to describe both neutral and charged excitations. For neutral excitations, the eDFT scheme is often termed Gross-Oliveria-Kohn DFT (GOK-DFT) as it is built on the variational principle introduced by the same authors [124, 125]. Senjean and Fromager then proposed an ensemble formalism, which they termed 'N-centered eDFT' that can describe both neutral and charged excitations [126]. These ensemble formulations offer several advantages over time-dependent approaches; most obviously the significantly less computationally demanding nature of these methods. In addition, eDFT can in theoretically model any kind of electronic excitation, some of which are not described by TD-DFT [127]. A more detailed review of the developments within eDFT can be found in Ref. [128].

In this chapter, we will recycle ideas proposed in the original fractional electron number extension of DFT [104], by constructing an OEP formulation for systems with a fractional average number of electrons.

6.1.2 The derivative discontinuity

One of the most common criticisms of many (semi-)local density functional approximations in DFT is their inability to calculate accurate fundamental gaps, with errors of around 40–50% frequently reported in the literature [120, 121]. A significant fraction of this large error is that these functionals lack a description of a non-analytical property of the exact KS potential known as the exchange-

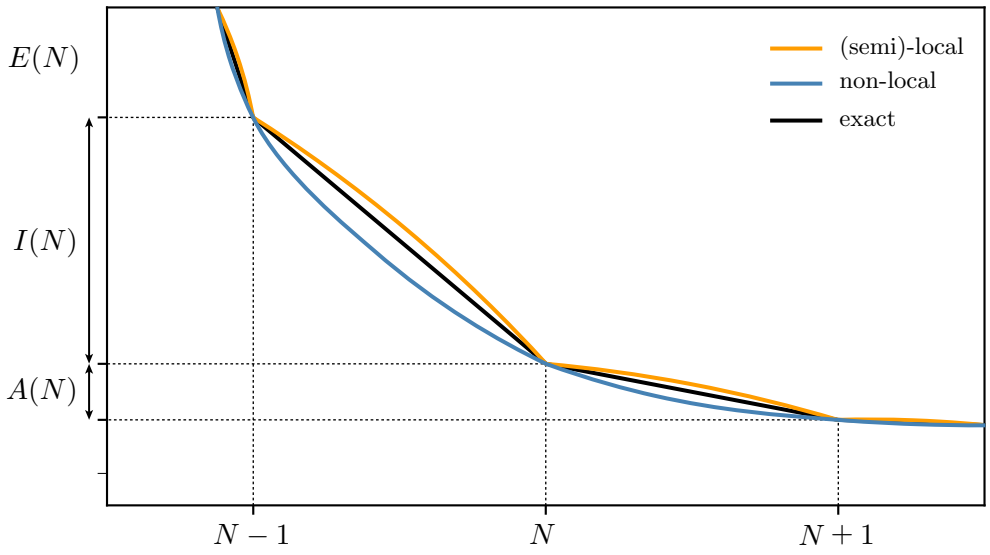


Figure 6.1: A schematic diagram (not actual calculations) of ensemble total energy curves obtained for the exact functional, a (semi-)local functional, and a non-local functional.

correlation derivative discontinuity, Δ_{xc}.

To understand the origin of this property, consider an open electronic system connected to a reservoir of electrons [104]. For this system the total energy varies linearly between integer numbers of electrons, however, its derivative exactly at an integer is discontinuous. The size of this discontinuity in the gradient of the total energy at N electrons defines the fundamental gap of the N electron system.

For the KS system, it is known that the HOMO eigenvalue has a direct physical meaning; minus the ionisation energy of the interacting system, i.e. $I^{(N)} = -\epsilon_{\text{ho}}^{(N)}$. The rest of the KS eigenvalues do not however correspond directly to an observable quantity of the interacting system; although, Janak's theorem [129] does prove that each of these eigenvalues is discontinuous at an integer number of electrons. Therefore, for a system with $N - \omega$ electrons when $\omega \to 0$, the HOMO eigenvalue is simply:

$$\lim_{\omega \to 0} \epsilon_{\text{ho}}^{(N-\omega)} = -I^{(N)}, \tag{6.1}$$

whereas a system consisting of $N + \omega$ in the same limit electrons possess a HOMO

eigenvalue of:

$$\lim_{\omega\to 0} \epsilon_{\mathrm{ho}}^{(N+\omega)} = -I^{(N+1)} = -A^{(N)}. \tag{6.2}$$

As the KS potentials for the $N-\omega$ and $N+\delta$ electron systems must yield the two above ionisation energies respectively, the exchange-correlation potential must shift by a constant (as the rest of the total energy functional is unaffected in this limit) as the number of electrons in the system exactly surpasses N electrons:

$$\Delta_{\mathrm{xc}} = \lim_{\omega\to 0} \left\{ v_{\mathrm{xc}}^{(N+\omega)}(\mathbf{r}) - v_{\mathrm{xc}}^{(N-\omega)}(\mathbf{r}) \right\}, \tag{6.3}$$

where Δ_{xc} is strictly a function of r, but in this limit of $\omega \to 0$, Δ_{xc} becomes a constant over all space.

From the above relations, the xc derivative discontinuity can also be expressed in terms of the ionisation energy, electron affinity, and the KS HOMO-LUMO gap for the N electron system:

$$\begin{aligned} \Delta_{\mathrm{xc}} &= \lim_{\omega\to 0} \left\{ I^{(N)} + \epsilon_N^{(N-\omega)} - \left[A^{(N)} + \epsilon_N^{(N+\omega)} \right] \right\} \\ &= I^{(N)} - A^{(N)} - \left[\epsilon_{\mathrm{lu}}^{(N)} - \epsilon_{\mathrm{ho}}^{(N)} \right]. \end{aligned} \tag{6.4}$$

Since $I^{(N)} = \epsilon_{\mathrm{ho}}^{(N)}$, it follows that that electron affinity is given by:

$$A^{(N)} = -\left[\epsilon_{\mathrm{lu}}^{(N)} + \Delta_{\mathrm{xc}} \right]. \tag{6.5}$$

Therefore, if it is possible to calculate the two potentials $v_{\mathrm{xc}}^{(N+\omega)}(\mathbf{r})$ and $v_{\mathrm{xc}}^{(N-\omega)}(\mathbf{r})$ in the region of $\omega \to 0$ around an integer number of N electrons, then one would be able to calculate directly of the derivative discontinuity for a given density functional approximation. In the following section, an OEP equation will be derived for an ensemble electronic system with a fractional average electron number which allows for the direct calculation of these two potentials.

6.1.3 The ensemble OEP equation

To derive an OEP equation that solves for the local potential that describes an ensemble system with fractional electron number between N and $N+1$, the starting

point is the definition of the ensemble total energy. For this type of ensemble, the total energy functional which interpolates between a system with N electrons at $\omega = 0$ and a system with $N+1$ electrons at $\omega = 1$, is given by:

$$\begin{aligned} E^{(N+\omega)}[\{\phi_i[\rho^{(N+\omega)}]\}] &= (1-\omega)E^{(N)}[\{\phi_i[\rho^{(N)}]\}] \\ &\quad + \omega E^{(N+1)}[\{\phi_i[\rho^{(N+1)}]\}];\ 0 \leq \omega < 1, \end{aligned} \tag{6.6}$$

where for $\nu \in \{N, N+1\}$,

$$\begin{aligned} E^{(\nu)}[\{\phi_i[\rho^{(\nu)}]\}] &= \sum_{i=1}^{\nu} \left\langle \phi_i[\rho^{(\nu)}] \right| t + v_{\text{en}} \left| \phi_i[\rho^{(\nu)}] \right\rangle + E_{\text{xc}}^{(\nu)}[\{\phi_i[\rho^{(\nu)}]\}] \\ &\quad + \frac{1}{2} \iint \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \frac{\rho^{(\nu)}(\mathbf{r})\rho^{(\nu)}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \end{aligned} \tag{6.7}$$

where t is the usual single particle kinetic energy operator and $v_{\text{en}}(\mathbf{r})$ is the electron-nuclear potential.

The ensemble density, $\rho^{(N+\omega)}(\mathbf{r})$, is given by a weighted sum of the N and $N+1$ electron subsystem densities:

$$\rho^{(N+\omega)}(\mathbf{r}) = (1-\omega)\rho^{(N)}(\mathbf{r}) + \omega\rho^{(N+1)}(\mathbf{r}), \tag{6.8}$$

where the ground state densities of the N and $N+1$ electron subsystems are constructed as:

$$\rho^{(N)}(\mathbf{r}) = \sum_{i=1}^{N} \rho_i^{(N)}(\mathbf{r}) = \sum_{i=1}^{N} \left| \phi_i[\rho^{(N)}](\mathbf{r}) \right|^2, \tag{6.9}$$

$$\rho^{(N+1)}(\mathbf{r}) = \sum_{i=1}^{N+1} \rho_i^{(N+1)}(\mathbf{r}) = \sum_{i=1}^{N+1} \left| \phi_i[\rho^{(N+1)}](\mathbf{r}) \right|^2. \tag{6.10}$$

It is important to note that the set of orbitals $\{\phi_i[\rho^{(N)}]\}$ and $\{\phi_i[\rho^{(N+1)}]\}$ are the occupied KS orbitals of the N and $N+1$ electron systems, respectively. The ensemble system possesses its own set of KS orbitals, $\{\phi_i[\rho^{(N+\omega)}]\}$, which are in general linear combinations of the KS orbitals for the two subsystems.

The ensemble density can therefore be written either in terms of the ensemble KS orbitals as:

$$\rho^{(N+\omega)}(\mathbf{r}) = \sum_{i=1}^{N} \left| \phi_i[\rho^{(N+\omega)}](\mathbf{r}) \right|^2 + \omega \left| \phi_{N+1}[\rho^{(N+\omega)}](\mathbf{r}) \right|^2, \tag{6.11}$$

or in terms of the KS orbitals of the subsystems:

$$\rho^{(N+\omega)}(\mathbf{r}) = \sum_{i=1}^{N}\left[(1-\omega)\left|\phi_i[\rho^{(N)}](\mathbf{r})\right|^2 + \omega\left|\phi_i[\rho^{(N+1)}](\mathbf{r})\right|^2\right] + \omega\left|\phi_{N+1}[\rho^{(N+1)}](\mathbf{r})\right|^2. \tag{6.12}$$

These ensemble orbitals satisfy their own set of KS equations:

$$\left[t + v_{\text{en}}(\mathbf{r}) + v_{\text{Hxc}}[\rho^{(N+\omega)}](\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}), \tag{6.13}$$

For conciseness, from this point on the explicit functional dependence on the density for energy and potential terms will be dropped in favour of a single superscript indicating the number of electrons.

Using the above definitions for the densities, the total ensemble energy functional can be re-expressed as:

$$\begin{aligned} E^{(N+\omega)} = {} & \sum_{i=1}^{N}\langle\phi_i|\, t + v_{\text{en}}\,|\phi_i\rangle + \omega\,\langle\phi_{N+1}|\, t + v_{\text{en}}\,|\phi_{N+1}\rangle \\ & + \frac{1}{2}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\frac{1}{|\mathbf{r}-\mathbf{r}'|}\Big\{\rho^{(N)}(\mathbf{r})\rho^{(N)}(\mathbf{r}') \\ & + \omega^2\rho_{N+1}(\mathbf{r})\rho_{N+1}(\mathbf{r}') + 2\omega\rho^{(N)}(\mathbf{r})\rho_{N+1}(\mathbf{r}')\Big\} \\ & + (1-\omega)E_{\text{xc}}^{(N)} + \omega E_{\text{xc}}^{(N+1)}, \end{aligned} \tag{6.14}$$

where the ensemble kinetic, Hartree, and xc functionals are defined respectively as:

$$T_s^{(N+\omega)} = \sum_{i=1}^{N}\langle\phi_i|\, t\,|\phi_i\rangle + \omega\,\langle\phi_{N+1}|\, t\,|\phi_{N+1}\rangle\,, \tag{6.15}$$

$$\begin{aligned} E_{\text{H}}^{(N+\omega)} = {} & \frac{1}{2}\iint \mathrm{d}\mathbf{r}\,\mathrm{d}\mathbf{r}'\frac{1}{|\mathbf{r}-\mathbf{r}'|}\Big\{\rho^{(N)}(\mathbf{r})\rho^{(N)}(\mathbf{r}') \\ & + \omega^2\rho_{N+1}(\mathbf{r})\rho_{N+1}(\mathbf{r}') \\ & + 2\omega\rho^{(N)}(\mathbf{r})\rho_{N+1}(\mathbf{r}')\Big\}, \end{aligned} \tag{6.16}$$

$$E_{\text{xc}}^{(N+\omega)} = (1-\omega)E_{\text{xc}}^{(N)} + \omega E_{\text{xc}}^{(N+1)}. \tag{6.17}$$

As the ensemble Hxc energy is an implicit functional of the ensemble density, the functional derivative must be taken using the OEP method. Keeping ω fixed, for

any value between 0 and 1, the total ensemble energy can be minimised indirectly by searching for the effective local potential, whose orbitals satisfy Equations 6.13 and minimise the total energy of the system. At the minimum, that potential is equal to $v_{\mathrm{Hxc}}^{(N+\omega)}(\mathbf{r})$, given by the functional derivative of the ensemble Hxc energy with respect to the ensemble density:

$$v_{\mathrm{Hxc}}[\rho^{(N+\omega)}](\mathbf{r}) = \left.\frac{\delta E_{\mathrm{Hxc}}^{(N+\omega)}}{\delta\rho(\mathbf{r})}\right|_{\rho^{(N+\omega)}}. \tag{6.18}$$

During the above variation the particle number does not change; it is always fixed once the ensemble weights have been chosen. When this potential is varied, $v_{\mathrm{Hxc}}^{(N+\omega)} \rightarrow v_{\mathrm{Hxc}}^{(N+\omega)} + \epsilon u$, it induces a variation in the ensemble KS orbitals:

$$|\phi_i'\rangle = \epsilon {\sum_n}' \frac{\langle\phi_i|\, u\, |\phi_n\rangle\, \langle\phi_n|}{\epsilon_i - \epsilon_n}. \tag{6.19}$$

This change in the orbitals then induces a corresponding change in the total ensemble energy. Substituting the perturbation to the orbitals into the form of the total ensemble energy in Equation 6.14 yields:

$$\begin{aligned}
\Delta E^{(N+\omega)} \;=\; & \epsilon \sum_{i=1}^{N} \frac{\langle\phi_{N+1}|\, u\, |\phi_i\rangle}{\epsilon_i - \epsilon_{N+1}} \Big\{ \langle\phi_i|\, t + v_{\mathrm{en}} + v_{\mathrm{H}}^{(N+\omega)} \\
& \qquad + (1-\omega) v_{\mathrm{xc}}^{(N)} + \omega v_{\mathrm{xc}}^{(N+1)}\, |\phi_{N+1}\rangle \Big\} \\
& + \epsilon \sum_{i=1}^{N} \sum_{a>N+1}^{\infty} \frac{\langle\phi_a|\, u\, |\phi_i\rangle}{\epsilon_i - \epsilon_a} \Big\{ \langle\phi_i|\, t + v_{\mathrm{en}} + v_{\mathrm{H}}^{(N+\omega)} \\
& \qquad + (1-\omega) v_{\mathrm{xc}}^{(N)} + \omega v_{\mathrm{xc}}^{(N+1)}\, |\phi_a\rangle \Big\} \\
& - \epsilon\, \omega \sum_{i=1}^{N} \frac{\langle\phi_{N+1}|\, u\, |\phi_i\rangle}{\epsilon_i - \epsilon_{N+1}} \Big\{ \langle\phi_i|\, t + v_{\mathrm{en}} + v_{\mathrm{H}}^{(N+1)} \\
& \qquad + v_{\mathrm{xc}}^{(N+1)}\, |\phi_{N+1}\rangle \Big\} \\
& + \epsilon\, \omega \sum_{a>N+1}^{\infty} \frac{\langle\phi_a|\, u\, |\phi_{N+1}\rangle}{\epsilon_{N+1} - \epsilon_a} \Big\{ \langle\phi_{N+1}|\, t + v_{\mathrm{en}} + v_{\mathrm{H}}^{(N+1)} \\
& \qquad + v_{\mathrm{xc}}^{(N+1)}\, |\phi_i\rangle \Big\} + \text{ c.c.}
\end{aligned} \tag{6.20}$$

At the minimum, the functional derivative ensemble energy functional must vanish. For the expression for the change in the ensemble energy above, the functional

derivative is given by:

$$
\begin{aligned}
&(1-\omega)\sum_{i=1}^{N}\frac{\langle\phi_{N+1}|\, v_{\mathrm{Hxc}}^{(N)}-v_{\mathrm{Hxc}}^{(N+\omega)}\,|\phi_i\rangle}{\epsilon_i-\epsilon_{N+1}}\phi_i^*(\mathbf{r})\phi_{N+1}(\mathbf{r})\\
&+\omega\sum_{a>N+1}^{\infty}\frac{\langle\phi_a|\, v_{\mathrm{Hxc}}^{(N+1)}-v_{\mathrm{Hxc}}^{(N+\omega)}\,|\phi_{N+1}\rangle}{\epsilon_{N+1}-\epsilon_a}\phi_{N+1}^*(\mathbf{r})\phi_a(\mathbf{r})\\
&+\sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\left\{\frac{\langle\phi_a|\, v_{\mathrm{H}}^{(N+\omega)}+(1-\omega)v_{\mathrm{xc}}^{(N)}+\omega v_{\mathrm{xc}}^{(N+1)}-v_{\mathrm{Hxc}}^{(N+\omega)}\,|\phi_i\rangle}{\epsilon_i-\epsilon_a}\right.\\
&\qquad\left.\times\,\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\right\}+\text{ c.c.}=\frac{\delta E^{(N+\omega)}}{\delta v(\mathbf{r})}=0. \qquad (6.21)
\end{aligned}
$$

Moving all of the terms that depend on the ensemble potential $v_{\mathrm{Hxc}}^{(N+\omega)}$ to the left-hand side of the above equation yields:

$$
\begin{aligned}
&\int \mathrm{d}\mathbf{r}'\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}')v_{\mathrm{Hxc}}^{(N+\omega)}(\mathbf{r}')\\
&\quad=\ (1-\omega)\sum_{i=1}^{N}\int\mathrm{d}\mathbf{r}'\frac{\phi_i^*(\mathbf{r})\phi_{N+1}(\mathbf{r})\phi_i(\mathbf{r}')\phi_{N+1}^*(\mathbf{r}')}{\epsilon_i-\epsilon_{N+1}}v_{\mathrm{Hxc}}^{(N)}(\mathbf{r}')\\
&\qquad+\omega\sum_{a>N+1}^{\infty}\int\mathrm{d}\mathbf{r}'\frac{\phi_{N+1}^*(\mathbf{r})\phi_a(\mathbf{r})\phi_{N+1}(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_{N+1}-\epsilon_a}v_{\mathrm{Hxc}}^{(N+1)}(\mathbf{r}')\\
&\qquad+\sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\int\mathrm{d}\mathbf{r}'\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}\\
&\qquad\quad\times\left\{v_{\mathrm{H}}^{(N+\omega)}+(1-\omega)v_{\mathrm{xc}}^{(N)}+\omega v_{\mathrm{xc}}^{(N+1)}\right\}+\text{ c.c.} \qquad (6.22)
\end{aligned}
$$

where the ensemble response function is defined as:

$$
\begin{aligned}
\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}') \;=\;& (1-\omega)\sum_{i=1}^{N}\frac{\phi_i^*(\mathbf{r})\phi_{N+1}(\mathbf{r})\phi_i(\mathbf{r}')\phi_{N+1}^*(\mathbf{r}')}{\epsilon_i-\epsilon_{N+1}}\\
&+\omega\sum_{a>N+1}^{\infty}\frac{\phi_{N+1}^*(\mathbf{r})\phi_a(\mathbf{r})\phi_{N+1}(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_{N+1}-\epsilon_a}\\
&+\sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}+\text{c.c.} \qquad (6.23)
\end{aligned}
$$

As the ensemble Hartree potential is given by the ensemble weighted sum of the N and $N+1$ subsystem Hartree potentials:

$$
v_{\mathrm{H}}^{(N+\omega)}(\mathbf{r})=(1-\omega)v_{\mathrm{H}}^{(N)}(\mathbf{r})+\omega v_{\mathrm{H}}^{(N+1)}(\mathbf{r}), \qquad (6.24)
$$

where:

$$
v_{\mathrm{H}}^{(\nu)}(\mathbf{r})=\int\mathrm{d}\mathbf{r}'\,\frac{\rho^{(\nu)}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \qquad (6.25)
$$

the final term on the right-hand side of Equation 6.22 can be condensed into Hxc potentials and written as two terms, each dependent on one of the two subsystems:

$$\begin{aligned}
&\int \mathrm{d}\mathbf{r}'\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}')v_{\mathrm{Hxc}}^{(N+\omega)}(\mathbf{r}') \\
&\quad = (1-\omega)\int \mathrm{d}\mathbf{r}'\Bigg[\sum_{i=1}^{N}\frac{\phi_i^*(\mathbf{r})\phi_{N+1}(\mathbf{r})\phi_i(\mathbf{r}')\phi_{N+1}^*(\mathbf{r}')}{\epsilon_i-\epsilon_{N+1}} \\
&\qquad + \sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}\Bigg]v_{\mathrm{Hxc}}^{(N)}(\mathbf{r}') \\
&\quad + \omega\int \mathrm{d}\mathbf{r}'\Bigg[\sum_{a>N+1}^{\infty}\frac{\phi_{N+1}^*(\mathbf{r})\phi_a(\mathbf{r})\phi_{N+1}(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_{N+1}-\epsilon_a} \\
&\qquad + \sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}\Bigg]v_{\mathrm{Hxc}}^{(N+1)}(\mathbf{r}') + \text{ c.c.}
\end{aligned} \tag{6.26}$$

The form of this ensemble response function in Equation 6.23 can be refactored into a far more intuitive form. To do so, we first split the third term in the response function into two terms, each with a weighting of $1-\omega$ and ω respectively:

$$\sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\{\dots\} = (1-\omega)\sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\{\dots\} + \omega\sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\{\dots\}. \tag{6.27}$$

Factoring out these weightings then allows the response function to be condensed into only two terms:

$$\begin{aligned}
\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}') &= (1-\omega)\Bigg[\sum_{i=1}^{N}\frac{\phi_i^*(\mathbf{r})\phi_{N+1}(\mathbf{r})\phi_i(\mathbf{r}')\phi_{N+1}^*(\mathbf{r}')}{\epsilon_i-\epsilon_{N+1}} \\
&\qquad + \sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}\Bigg] \\
&\quad + \omega\Bigg[\sum_{a>N+1}^{\infty}\frac{\phi_{N+1}^*(\mathbf{r})\phi_a(\mathbf{r})\phi_{N+1}(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_{N+1}-\epsilon_a} \\
&\qquad + \sum_{i=1}^{N}\sum_{a>N+1}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}\Bigg] + \text{ c.c.}
\end{aligned} \tag{6.28}$$

Next, the sums within these terms can be combined. For the first term, the sum over i shifts the lower index on the sum over a from $N+2$ to $N+1$. For the second, the sum over a shifts the upper index on the sum over i from N to $N+1$.

Together, these changes allow the ensemble response function to be written as:

$$\begin{aligned}\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}') &= (1-\omega)\sum_{i=1}^{N}\sum_{a=N+1}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a} \\ &\quad+\omega\sum_{i=1}^{N+1}\sum_{a>N+1}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}+\text{ c.c.}\end{aligned} \tag{6.29}$$

Recalling the original definition of the response function in Equation 3.59:

$$\chi^{(\nu)}(\mathbf{r},\mathbf{r}')=\sum_{i}^{\nu}\sum_{a>\nu}^{\infty}\frac{\phi_i^*(\mathbf{r})\phi_a(\mathbf{r})\phi_i(\mathbf{r}')\phi_a^*(\mathbf{r}')}{\epsilon_i-\epsilon_a}+\text{c.c.};\ \nu\in\{N,N+1\}, \tag{6.30}$$

we can identify each of the two terms in the above expression as none other than the response functions for each of the two subsystems, $\chi^{(N)}$ and $\chi^{(N+1)}$, respectively.

The total ensemble function can therefore simply be written as the ensemble weighted sum of the response functions for each of the subsystems:

$$\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}')=(1-\omega)\chi^{(N)}(\mathbf{r},\mathbf{r}')+\omega\chi^{(N+1)}(\mathbf{r},\mathbf{r}'). \tag{6.31}$$

In a similar manner to the ensemble response function, the right-hand side of Equation 6.26, which we will define as $b_{\text{Hxc}}^{(N+\omega)}(\mathbf{r})$, can also be expressed in terms of the subsystem response functions $\chi^{(N)}$ and $\chi^{(N+1)}$ by performing the same tricks on the indices on the summations:

$$\begin{aligned}b_{\text{xc}}^{(N+\omega)}(\mathbf{r}) &= (1-\omega)\int\mathrm{d}\mathbf{r}'\chi^{(N)}(\mathbf{r},\mathbf{r}')v_{\text{Hxc}}^{(N)}(\mathbf{r}') \\ &\quad+\omega\int\mathrm{d}\mathbf{r}'\chi^{(N+1)}(\mathbf{r},\mathbf{r}')v_{\text{Hxc}}^{(N+1)}(\mathbf{r}'),\end{aligned} \tag{6.32}$$

where the potentials $v_{\text{Hxc}}^{(N)}$ and $v_{\text{Hxc}}^{(N+1)}$ are known from the integer N and $N+1$ electrons systems.

Finally, putting everything together, the OEP equation that determines the exact ensemble potential, $v_{\text{Hxc}}^{(N+\omega)}$, is:

$$\int\mathrm{d}\mathbf{r}'\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}')v_{\text{Hxc}}^{(N+\omega)}(\mathbf{r}')=b_{\text{Hxc}}^{(N+\omega)}(\mathbf{r}), \tag{6.33}$$

with the ensemble xc potential obtained directly by inverting the above equation:

$$v_{\text{Hxc}}^{(N+\omega)}(\mathbf{r})=\int\mathrm{d}\mathbf{r}'\left(\chi^{(N+\omega)}(\mathbf{r},\mathbf{r}')\right)^{-1}b_{\text{Hxc}}^{(N+\omega)}(\mathbf{r}'). \tag{6.34}$$

Equation 6.33, and its corresponding inverted form in Equation 6.34, is an important result, defining for the first time the formally exact KS DFT equations for an ensemble system with a fractional number of electrons. Using this equation, the exact ensemble potential can be constructed in terms of the ground-state v_{Hxc} potentials for an integer N and $N+1$ electrons without incurring any approximations (note again that the ensemble potential is defined for fixed ω up to a constant, which is chosen such that $v_{\mathrm{Hxc}}^{(N+\omega)}(r) \to 0$ for $r \to \infty$).

To use this OEP formulation in practice, density functional approximations must, however, be invoked in order to obtain the xc potentials $v_{\mathrm{xc}}^{(N)}$ and $v_{\mathrm{xc}}^{(N+1)}$. Therefore, the accuracy of calculated properties is limited by the underlying approximation that is employed.

6.1.4 Numerical implementation

In order to solve numerically the ensemble OEP equation 6.33 derived above, it must first be implemented in our code for finite basis sets. To begin, the ensemble xc potential is expanded in terms of a screening density, in line with the previous implementations of the OEP method for finite basis set calculations used in the constrained and implicit methods. This screening density is then expanded in terms of an auxiliary basis set $\{\theta_k\}$:

$$v_{\mathrm{Hxc}}^{(N+\omega)}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \, \frac{\rho_{\mathrm{scr}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}; \quad \rho_{\mathrm{scr}}(\mathbf{r}) = \sum_{k=1}^{n_{\mathrm{aux}}} \rho_k^s \theta_k(\mathbf{r}). \tag{6.35}$$

Inserting this expansion into the OEP equation gives:

$$\sum_{k=1}^{n_{\mathrm{aux}}} \rho_k^s \int \chi^{(N+\omega)}(\mathbf{r},\mathbf{r}') \tilde{\theta}_l(\mathbf{r}') = b_{\mathrm{Hxc}}^{(N+\omega)}(\mathbf{r}), \tag{6.36}$$

where $\tilde{\theta_k}(\mathbf{r}') = \int \mathrm{d}\mathbf{x} \; \theta_k(\mathbf{x})/|\mathbf{r}'-\mathbf{x}|$.

Multiplying both sides by $\tilde{\theta_k}(\mathbf{r})$ and integrating over $\mathrm{d}\mathbf{r}$ yields a matrix equation for the coefficients $\{\rho_k^s\}$:

$$\rho_k^s = \sum_{l=1}^{n_{\mathrm{aux}}} \left(A^{(N+\omega)}\right)_{kl}^{-1} b_{\mathrm{Hxc},\,l}^{(N+\omega)}, \tag{6.37}$$

where,

$$A_{kl}^{(N+\omega)} = \left\langle \tilde{\theta}_k \right| \chi^{(N+\omega)} \left| \tilde{\theta}_l \right\rangle, \; b_{\mathrm{Hxc},\,l}^{(N+\omega)} = \left\langle b_{\mathrm{Hxc}}^{(N+\omega)} \middle| \tilde{\theta}_l \right\rangle. \tag{6.38}$$

Once again, it is necessary to introduce a constraint on the screening density, where $Q_{\mathrm{scr}} = \int \mathrm{d}\mathbf{r}\rho_{\mathrm{scr}}(\mathbf{r})$. For a regular OEP calculation for the Hxc potential of an N electron system, each electron should be repelled by $N-1$ electrons; as electrons should not experience self-interactions. However, the choice of screening charge for the ensemble of N and $N+1$ electron systems is not as obvious [102]. Considering the behaviour of the ensemble OEP potential in the asymptotic regions, it is known that the decay of the density depends only on the eigenvalue of the HOMO:

$$\lim_{r\to\infty} \rho(\mathbf{r}) \propto e^{-2\sqrt{-2\epsilon_{\mathrm{ho}}}|\mathbf{r}|}. \tag{6.39}$$

So, in the case of the ensemble system, as soon as an infinitesimal amount of an electron, δ, is added such that the total number of electrons surpasses the integer number of electrons, N, the effective electrons 'felt' by each electron increases discontinuously from $N-1$ to N; this is due to the partially occupied HOMO of the $N+1$ electron system now being also the HOMO of the $N+\delta$ electron system.

To include this screening charge constraint, the matrix equation can be reformulated as a minimisation of an objective functional, $G[\{\rho_k^s\}]$, with respect to the expanded screening density coefficients:

$$G[\{\rho_k^s\}] = \frac{1}{2}\sum_{k=1}^{n_{\mathrm{aux}}}\sum_{l=1}^{n_{\mathrm{aux}}} \rho_k^s A_{kl}^{(N+\omega)} \rho_l^s - \sum_{l=1}^{n_{\mathrm{aux}}} \left(b_{\mathrm{Hxc},\,l}^{(N+\omega)} + \lambda X_l\right)\rho_l^s, \tag{6.40}$$

where $X_l = \int \mathrm{d}\mathbf{r}\; \theta_l(\mathbf{r})$ and λ is a Lagrange multiplier that enforces the constraint on the screening charge. The coefficients $\{\rho_k^s\}$ can then be determined by solving the following equation:

$$\rho_k^s = \sum_{l=1}^{n_{\mathrm{aux}}} \left(A^{(N+\omega)}\right)_{kl}^{-1} \left(b_{\mathrm{Hxc},\,l}^{(N+\omega)} - \lambda X_l\right). \tag{6.41}$$

To obtain the value of λ the above expression can be inserted into the definition of

the screening charge:

$$\begin{aligned} Q_{\text{scr}} &= \sum_{k=1}^{n_{\text{aux}}} \rho_k^s X_k = \sum_{k=1}^{n_{\text{aux}}} \left[\sum_{l=1}^{n_{\text{aux}}} \left(A^{(N+\omega)} \right)_{kl}^{-1} \left(b_{\text{Hxc},\,l}^{(N+\omega)} - \lambda X_l \right) \right] X_k \\ \implies \lambda &= \frac{\sum_{k=1}^{n_{\text{aux}}} \sum_{l=1}^{n_{\text{aux}}} X_k \left(A^{(N+\omega)} \right)_{kl}^{-1} b_{\text{Hxc},\,l}^{(N+\omega)} - Q_{\text{src}}}{\sum_{k=1}^{n_{\text{aux}}} \sum_{l=1}^{n_{\text{aux}}} X_k \left(A^{(N+\omega)} \right)_{kl}^{-1} X_l} \end{aligned} \tag{6.42}$$

The matrix $A_{kl}^{(N+\omega)}$ can be easily constructed from the subsystem response functions $\chi^{(N)}(\mathbf{r}, \mathbf{r}')$ and $\chi^{(N+1)}(\mathbf{r}, \mathbf{r}')$, by expanding them both on the orbital basis set. The right-hand side of the equation however requires some more care to calculate. Conveniently, expressions from the implicit method can be utilised here, as each of the terms can be expanded into the spin-up and spin-down channels for each subsystem:

$$\begin{aligned} b_{\text{Hxc}}^{(N+\omega)}(\mathbf{r}) &= \sum_{\sigma=\uparrow\downarrow} \int \mathrm{d}\mathbf{r}' \Big\{ \chi^{(N)^\sigma}(\mathbf{r}, \mathbf{r}') v_{\text{Hxc}}^{(N)^\sigma}(\mathbf{r}') \\ &\quad + \omega \chi^{(N+1)^\sigma}(\mathbf{r}, \mathbf{r}') v_{\text{Hxc}}^{(N+1)^\sigma}(\mathbf{r}') \Big\}. \end{aligned} \tag{6.43}$$

Therefore, the above expression allows the vector $b_{\text{Hxc},\,l}^{(N+\omega)}$ to be constructed solely from quantities recycled from the implicit OEP implementation.

To continue, the potentials $v_{\text{Hxc}}^{(N)^\sigma}$ and $v_{\text{Hxc}}^{(N+1)^\sigma}$ can each be split into the sum of a Hartree and an xc potential. As in previous OEP implementations discussed in this thesis, the individual xc potentials can be calculated via the LIBXC package [94]. Given these potentials on the grid, they can be transformed to the auxiliary basis:

$$\rho_{\text{xc},k}^{(\nu)^\sigma} = \sum_{k=1}^{n_{\text{aux}}} \left\langle \tilde{\theta_k} \middle| \theta_l \right\rangle^{-1} \left\langle \theta_l \middle| v_{\text{xc}}^{(\nu)^\sigma} \right\rangle; \quad \nu \in \{N, N+1\}. \tag{6.44}$$

Finally, the vector $b_{\text{Hxc},\,l}^{(N+\omega)}$ can then be expressed as:

$$b_{\text{Hxc},\,l}^{(N+\omega)} = \sum_{\nu} \sum_{k=1}^{n_{\text{aux}}} \omega_\nu \left\{ \rho_{\text{H},k}^{(\nu)} A_{kl}^{(\nu)} + \sum_{\sigma=\uparrow\downarrow} \rho_{\text{xc},k}^{(\nu)^\sigma} A_{kl}^{(\nu)^\sigma} \right\}, \tag{6.45}$$

where ω_ν are the respective weighting factors of $\omega_N = (1-\omega)$ and $\omega_{N+1} = \omega$.

At this point, the form of all quantities needed for the ensemble OEP implementation are known and can be implemented into a finite basis set code. Once again,

this work was incorporated into the HIPPO code, building on existing routines developed for the implicit OEP method discussed in the previous chapter. To avoid any spurious oscillations in the potentials, the complement method for χ in a finite orbital basis was again used in order to obtain well-behaved OEP solutions.

6.2 Results

6.2.1 Ensemble energies

One of the first quantities that can be investigated for this new OEP method is the total ensemble ground-state energy. Due to the form of the ensemble energy functional in Equation 6.6, the ensemble OEP should yield a linear variation in the ground-state total energy along the variation of $\omega = 0 \rightarrow 1$.

In Figure 6.2, the total ensemble energy is plotted, starting from the Be^{2+} ion up to the Be atom, for the ensemble OEP method (using the LDA functional), implicit OEP, and regular unpolarised LDA. This figure shows a very strong resemblance to the exact and (semi-)local curves from Figure 6.1. Both the implicit and unpolarised LDA vary continuously, i.e. there is no discontinuous change in their gradient at the exact point where the number of electrons in the ensemble surpasses $N = 3$ (i.e. at the Be^{+} ion).

On the other hand, the ensemble LDA curve does exhibit the desired discontinuous change in the gradient of the energy curve, and therefore possesses a non-zero derivative discontinuity. We also stress that although the ensemble energy is expected to interpolate linearly between these limits, this is the first time that this has been demonstrated using LDA for ensemble systems.

Focusing instead on the bottom subplot of Figure 6.1, which plots the difference relative to the ensemble energy, we observe that the implicit LDA energy at an integer number of electrons is equal to the ensemble energy. The unpolarised LDA curve does match at the endpoints where the number of electrons is even, but it

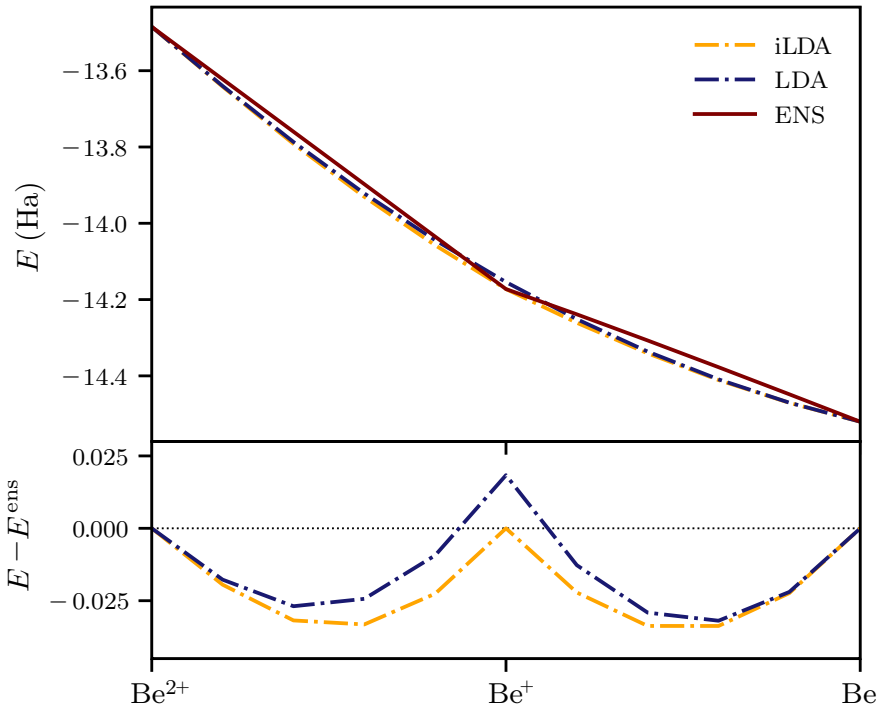


Figure 6.2: Total energy calculations ranging from the Be^{2+} ion to the Be atom for ensemble OEP method (red), compared to regular unpolarised LDA (blue) and ghost exchange corrected LDA (orange). All calculations use an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

is not equal to the ensemble curve at $N = 3$. This difference between the iLDA and the LDA curves is a clear example of the ghost exchange error in action, with the difference between these curves at $N = 3$ defining the value of this error for the Be^+ ion. As the ensemble OEP method is constructed as an extension of the implicit OEP, it directly incorporates the ghost exchange correction. We should also note that although it is not plotted (as it yields total energies that are very close to those obtained from iLDA), the spin-polarised LSDA functional also does not possess a non-zero derivative discontinuity in the total energy; only the new ensemble method has a non-zero change in the energy derivative at integer N.

As the ensemble energy exhibits a discontinuous change in its gradient, it is, therefore, possible to calculate the gradient for $N = 3 \pm \omega \to 0$ to estimate the value of the xc derivative discontinuity. Doing so would then enable a new value for the

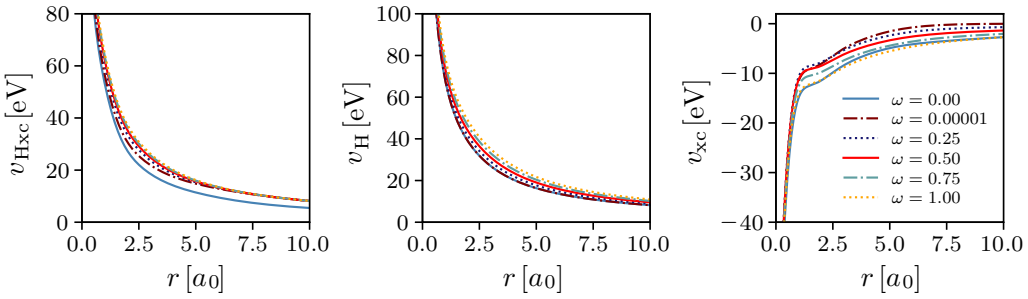


Figure 6.3: Potentials obtained using the ensemble LDA method for the Be^+ ion across a range of values of ω. The ensemble Hartree potential (middle) is subtracted from the Hxc potential (left) to obtain the ensemble xc potential (right). For each of these calculations, the orbital and auxiliary basis sets used were augmented cc-pVTZ uncontracted cc-pVTZ respectively.

electron affinity to be calculated using Equation 6.5. Results obtained using this method will be presented and discussed in § 6.2.3.

6.2.2 Ensemble potentials

As in the preceding two chapters, we should also take a closer look at the OEP potentials that are calculated via this new ensemble formulation, across the entire range of values of the ensemble weight, ω. In § 5.2.1, we investigated the convergence of the OEP potentials with respect to the basis sets and the value of α. For the systems studied in this thesis, we found that a value of $\alpha = 10^{-2}$ consistently yielded well-behaved and converged potentials, and so in the following figures we will fix $\alpha = 10^{-2}$. All results presented (unless otherwise stated) employ the LDA approximation in the construction of the exchange-correlation potentials.

Figure 6.3 highlights how the underlying potentials; v_{Hxc}, v_{H}, and v_{xc}, vary as the value of omega is increased from 0 to 1 for the Be^+ ion. From the left-hand subplot, we can clearly see that as a very small amount of an electron is added (here we show $\omega = 10^{-5}$) to the ion, the asymptotic tail of the Hxc potential shifts discontinuously. However, as we continue to increase $\omega \to 1$, the asymptotic tail exhibits no further variation. This is a direct consequence of the intuitive constraints that were placed on the screening density (see § 6.1.4 for a more detailed

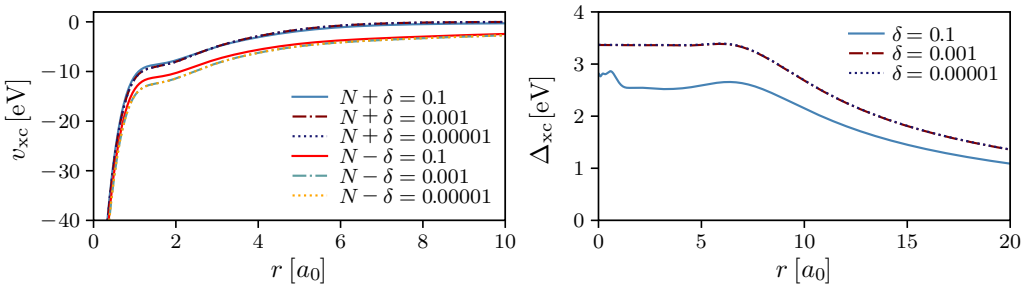


Figure 6.4: Calculated ensemble xc potentials for the Be^+ ion for increasingly smaller values of δ. On the right-hand side, the difference between the xc potentials for the $N+\delta$ and $N-\delta$ ensemble electron systems are plotted (Δ_{xc}). The basis sets were kept fixed with are calculations using an augmented cc-pVTZ orbital basis and an uncontracted cc-pVTZ auxiliary basis.

discussion). As an infinitesimal amount of an electron is added to the ion, it begins to occupy the LUMO of the $\omega = 0$ system. Therefore, as the decay of the electronic density is related to the highest occupied eigenvalue (see Equation 6.39) the screening charge must change discontinuously to compensate.

Studying the middle plot of Figure 6.3, it is clear that across the entire range of values of ω the ensemble Hartree potential varies linearly; there is no discontinuous shift present. We can therefore conclude that the shift in the Hxc potential is due to a shift in the xc potential. This is confirmed by inspecting the right-hand plot in Figure 6.3, where the xc potential does indeed change discontinuously on the addition of a very small fraction of an electron. Another observation that can be made from this figure, is that after the 'jump', the xc potential moves closer to the $\omega = 0$ result as $\omega \to 1$. If we again consider the value of the screening charge along this range, this should be expected. At $\omega = 0$, the Hxc screening charge is equal to $N - 1$, and the effective xc screening charge is equal to -1. As a small fraction of an electron is added, the Hxc screening charge jumps to N, and therefore the xc screening charge shifts to 0. Continuing to add a larger fraction of an electron until $\omega \to 1$, the Hxc screening charge remains unchanged. However, the xc screening charge will increase linearly until it reaches -1. This, therefore, explains why the xc potentials calculated at $\omega = 0$ and $\omega \to 1$ exhibit the same asymptotic behaviour.

This observed shift in the xc potential as an infinitesimal amount of an electron is worth investigating in more detail. From Equation 6.3, we know that the xc derivative discontinuity can be calculated from the difference between the two ensemble potentials with an infinitesimal amount of an electron, δ, added and removed from the N electron system. In the above discussion, we have already calculated the $N + \delta$ xc potential, therefore we also need to calculate the potential for $N - \delta$ electrons. This is very straightforward to accomplish using the ensemble method, as the $N - \delta$ xc potential is exactly the $N + (1 - \delta)$ potential for the ensemble of $N - 1$ and N electron subsystems.

In the left-hand plot of Figure 6.4, the ensemble xc potentials for the $N + \delta$ and $N - \delta$ ensembles are plotted for three increasingly smaller values of δ. As before, we observe a shift in the xc potentials as the fractional amount of an electron is added, however, the shift in the two potentials for fixed δ appears to converge as $\delta \to 0$. If we inspect the differences between the two in the right-hand figure, this becomes very apparent; with the change appearing to converge to a constant within the vicinity of the atom.

This shift is exactly the definition of the xc derivative discontinuity, although here we do not see a constant shift across all space. At first, it was assumed that this decay was an artifact of the (semi-)local density functional approximations employed in the ensemble OEP equation to construct the individual spin-xc potentials for the two subsystems. This however is not likely to be the only reason, as independent results obtained from the inversion of full configuration interaction (CI) densities for finite basis sets have recently been shown to yield similar behaviour; a constant shift in the xc potential that decays to zero far from the system [130]. Therefore, we can conclude that this decay is also a Gaussian basis set effect.

We can also investigate the convergence of this constant with respect to the auxiliary basis set, to ensure that we are not simply observing an artifact from using basis sets that are too small and that it does not vanish in the complete basis set limit. In Figure 6.5, the convergence of Δ_{xc} with respect to the size of the auxili-

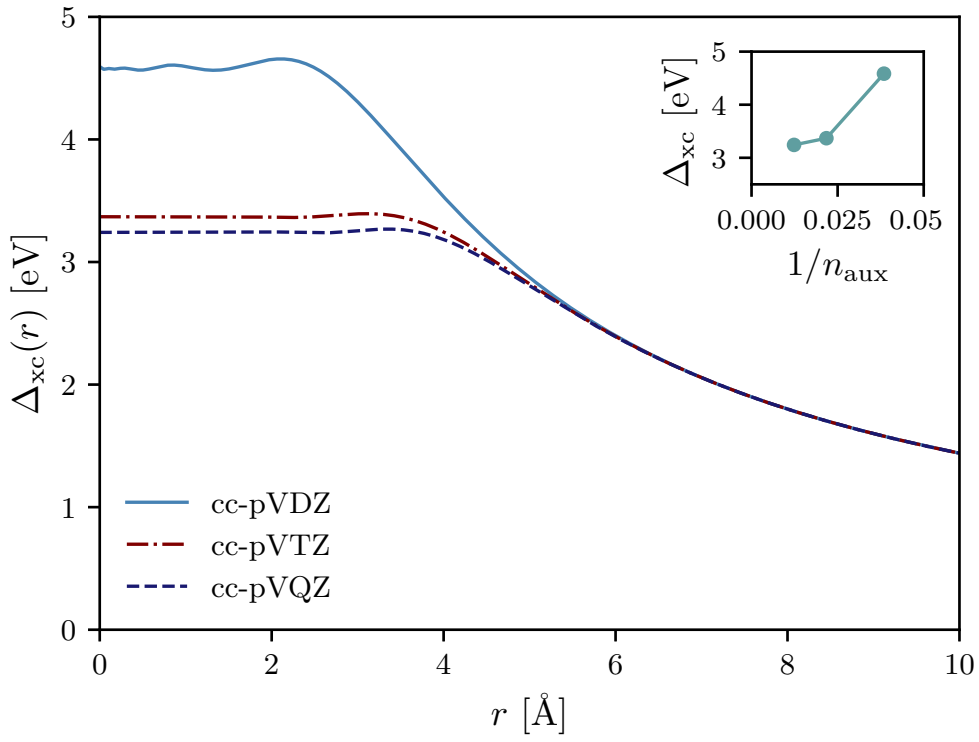


Figure 6.5: Convergence of the derivative discontinuity, $\Delta_{xc}(r)$, with respect to the size of the auxiliary basis set for the Be^+ ion, calculated via the difference between the two ensemble potentials $v_{xc}^{(N+\delta)}(r)$ and $v_{xc}^{(N-\delta)}(r)$ for $\delta = 10^{-5}$. For all calculations a fixed augmented cc-pVQZ orbital basis was used, with the auxiliary basis set varying from uncontracted cc-pVDZ to QZ.

ary basis set is plotted for a fixed augmented cc-pVQZ orbital basis, for a value of $\delta = 10^{-5}$. A similar result to the above is observed, with the constant appearing to converge to a flat plateau in the region of the atom, decaying off at large distances. The spatial extent of this plateau increases slightly with the size of the auxiliary basis, whereas its height decreases; although the drop in its magnitude is shown to converge well with the number of auxiliary basis set functions (see nested subplot).

So far, we have only studied the ensemble potentials obtained for the single atom Be^+ ion; although we can still observe the same behaviour in larger systems. Figure 6.6 plots the derivative discontinuity obtained from the difference in ensemble potentials for a molecule of Benzene (C_6H_6). Across the zx-plane, we observe a constant Δ_{xc} within the neighborhood of the molecule. A similar picture can also

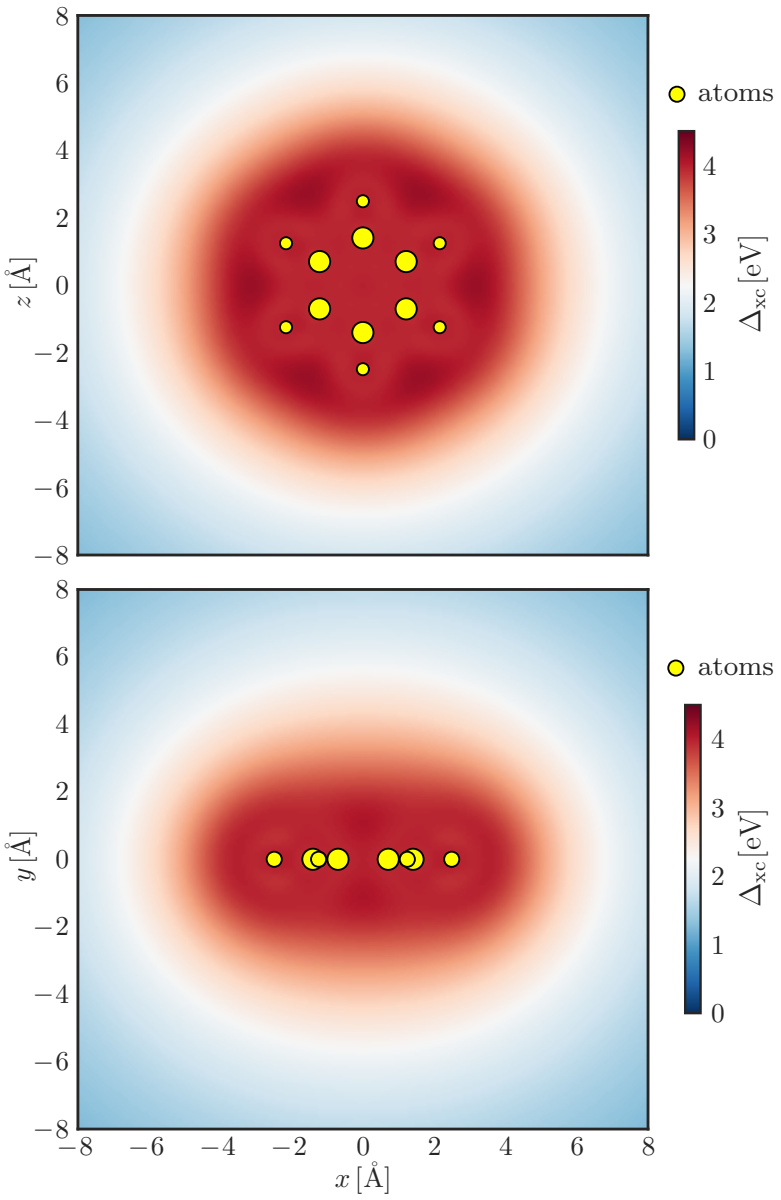


Figure 6.6: The calculated derivative discontinuity for Benzene (C_6H_6) via the difference of the ensemble potentials for $N \pm \delta$, $\delta = 10^{-5}$. The two calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

be obtained by plotting the variation in the plane perpendicular to the molecule, with the shift manifesting in all three real-space dimensions.

Convergence of the derivative discontinuity for molecules with respect to the auxiliary basis set can also be checked; with Figure 6.7 presenting the variation in the magnitude of Δ_{xc} for a molecule of H_2O for an increasing auxiliary basis set size.

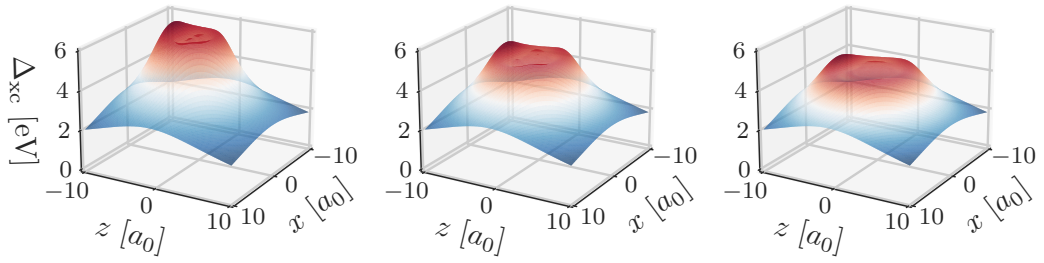


Figure 6.7: The calculated derivative discontinuity for H_2O via the difference of the ensemble potentials for $N \pm \delta$, $\delta = 10^{-5}$. These calculations were performed using an augmented cc-pVQZ orbital basis set and an uncontracted (left) cc-pVDZ, (middle) cc-pVTZ, and (right) cc-pVQZ auxiliary basis set.

Again, the value appears to plateau in the vicinity of the system with the change in magnitude decreasing as the basis size increases.

Now that it has been demonstrated that the ensemble OEP can yield a non-zero derivative discontinuity in the region of the system which does not vanish in the complete basis set limit, we can now investigate the magnitude of the Δ_{xc} calculated via this difference and compare it against values calculated from the difference in the gradient of the total ensemble energy as the number of electrons surpasses an integer.

6.2.3 Estimating the derivative discontinuity

As already mentioned, (semi-)local density functional approximations do not generally possess a non-zero derivative discontinuity. However, we have seen in the previous two sections that by solving the ensemble OEP equation with these functionals, we can in fact obtain a non-zero Δ_{xc} for any (semi-)local DFA.

First, we can calculate the total ensemble ground state energy for a system from $N-1$ to N electrons, the total ensemble energy from N to $N+1$ electrons, and directly calculate the change in the gradient of these total energy curves. The ensemble total varies linearly along the path of $w = 0 \rightarrow 1$ and changes discontinuously at exactly N electrons (see Figure 6.2). Calculating the difference

System	Exp	$\Delta \frac{\mathrm{d}E^{\mathrm{eLDA}}}{\mathrm{d}N}$	$\Delta_{\mathrm{xc}}^{\omega=0}$	$\Delta_{\mathrm{xc}}^{\omega=1}$
Li	4.77	4.95	2.54	5.14
Be^+	8.89	9.25	3.37	9.06
Na	4.59	4.81	2.65	5.20
F	14.02	17.41	8.18	14.26
OH	11.19	14.12	5.35	11.43
CN	9.74	10.68	5.20	10.20
NH_2	10.01	11.74	4.97	10.48
HS	8.11	9.71	4.81	8.59
ClO	8.61	10.05	5.33	9.24
CNO	8.15	9.94	4.90	9.13
SCN	7.15	8.13	4.34	7.61
|Avr. Err.|	—	1.41	3.96	0.46

Table 6.1: Values of the derivative discontinuity calculated for a set of small systems compared to experimental fundamental gaps taken from the Computational Chemistry Comparison and Benchmark Database (CCCBDB) [114]. The third column was computed by calculating the difference between the gradient of the ground state total ensemble energy at $N-\omega$ and $N+\omega$ electrons for $\omega \to 0$ using the LDA functional. The remaining two columns were calculated using the ensemble OEP method via the difference in the xc potentials. All calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

in the gradient at this point directly yields the value of Δ_{xc}.

Second, we can determine a value of the derivative discontinuity from the height of the step in the vicinity of the system that manifests in the difference between the two ensemble potentials at $N \pm \omega \to 0$ (as discussed in § 6.2.2), which we will denote here as $\Delta_{\mathrm{xc}}^{\omega=0}$.

In Table 6.1, we present results obtained for the derivative discontinuity for a small set of systems using the LDA functional. As these systems are open-shell and the ensemble OEP is a spin-restricted method (i.e. $\phi_i^{\uparrow} = \phi_i^{\downarrow}$), the value of Δ_{xc} yields the fundamental gap, $I - A$, of the system, allowing us to directly compare against experimental values from the CCCBDB [114]. For these examples, it is clear that the value of Δ_{xc} calculated from the change in the energy gradient is more accurate than the shift in the xc potential, with an average absolute error of 1.41 eV compared to 3.96 eV. A similar set of results are obtained using the PBE

System	Exp	$\Delta \frac{\mathrm{d}E^{\mathrm{ePBE}}}{\mathrm{d}N}$	$\Delta_{\mathrm{xc}}^{\omega=0}$	$\Delta_{\mathrm{xc}}^{\omega=1}$
Li	4.77	5.06	2.50	4.52
Be^+	8.89	9.49	3.38	10.31
Na	4.59	4.72	5.30	7.18
F	14.02	17.61	8.01	13.05
OH	11.19	14.35	5.27	10.64
CN	9.74	11.06	5.13	9.40
NH_2	10.01	11.96	4.88	7.85
HS	8.11	9.89	4.73	7.93
ClO	8.61	10.17	5.24	8.46
CNO	8.15	10.06	4.83	8.73
SCN	7.15	8.23	4.28	7.24
|Avr. Err.|	—	1.58	3.92	0.84

Table 6.2: Values of the derivative discontinuity calculated for a set of small systems compared to experimental fundamental gaps taken from the CCCBDB [114]. The third column was computed by calculating the difference between the gradient of the ground state total ensemble energy at $N - \omega$ and $N + \omega$ electrons for $\omega \to 0$ using the PBE functional. The remaining two columns were calculated using the ensemble OEP method via the difference in the xc potentials. All calculations were performed using an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

functional in Table 6.2, where the gradient difference method yields an average error of 1.58 eV compared to the average error of 3.92 eV determined from the shift in the ensemble potentials.

Clearly, the estimate obtained via the step that manifests in the region of the system is too small to account for the full value of the derivative discontinuity. We can gain a deeper understanding of why this may be the case by plotting the value of the HOMO eigenvalue for the ensemble as the number of electrons varies from N to $N+1$. In Figure 6.8, the value of the HOMO eigenvalue is plotted for the Be atom with a varying number of electrons. At $N = 3$, we observe a discontinuous change in the eigenvalue, which occurs as a consequence of the discontinuous change in the screening charge from $2 \to 3$. The magnitude of this shift is also approximately equal to the magnitude of the shift in the xc potentials as the number of electrons surpasses $N = 3$ and is, therefore, an alternative measure of $\Delta_{\mathrm{xc}}^{\omega=0}$. From Janak's theorem [129], we would have expected that throughout the range of $\omega = 0 \to 1$

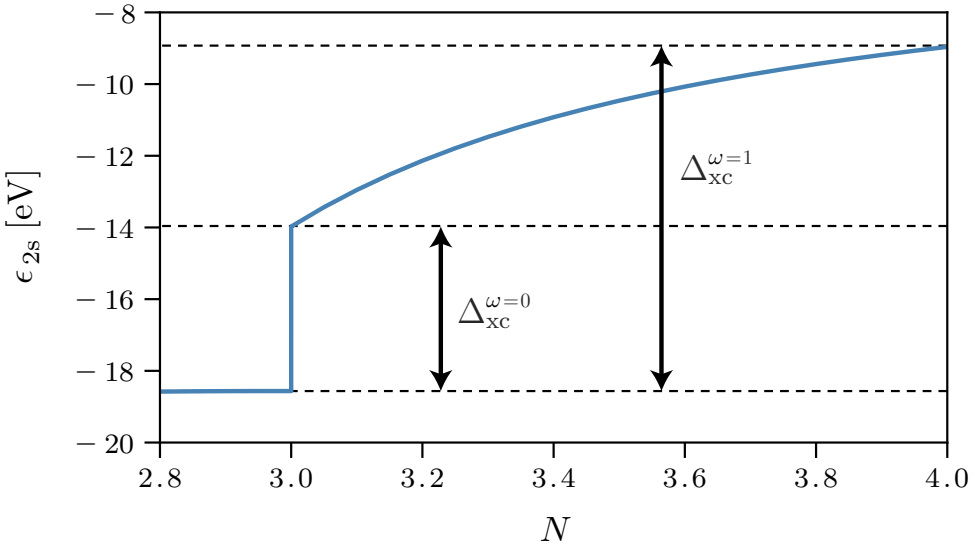


Figure 6.8: Variation of the value of the 2s eigenvalue for the Be atom obtained via the ensemble OEP method with respect to the average number of electrons in the ensemble system. All calculations used an augmented cc-pVTZ orbital basis and an uncontracted cc-pVTZ auxiliary basis.

the HOMO, eigenvalue would remain a constant, however, it is clear from this figure, that this is not the case for the ensemble OEP method. As the number of electrons increases from $3 \rightarrow 4$, the HOMO eigenvalue also increases. The value of the HOMO eigenvalue at $N = 4$ (i.e. $\omega = 1$), corresponds to the HOMO eigenvalue of the $N+1$ electron system which we had used previously as an accurate measure of the electron affinity in § 4.2.5. By defining the total variation of the eigenvalue as $\Delta_{xc}^{\omega=1}$, we obtain a third approximation for the fundamental gap, which yielded an average error of 0.46 eV and 0.84 eV for LDA and PBE respectively, a more accurate estimate for than the two alternative methods.

Using these results, we argue that although the step in the potentials that occurs as the number of electrons surpasses an integer is too small, the discontinuous shift in the screening charge does in fact contribute a large fraction to the total derivative discontinuity. This shift in the screening charge is a physically intuitive constraint, that can systematically be applied to any (semi-)local DFA to yield a non-zero Δ_{xc} with relative ease. The remaining fraction of the derivative discontinuity in this

method comes from the additional variation in the eigenvalue, which is expected to remain constant.

6.3 Summary

In this chapter, we formulated and derived for the first time the exact ensemble KS equations for systems with a fractional number of electrons via a new OEP scheme. This formulation extended the ideas introduced in implicit OEP, by constructing each subsystem in a ghost exchange free manner. The ensemble OEP exhibits several desired features of the exact KS potential that are lacking from (semi-)local density functional approximations, in particular a non-zero xc derivative discontinuity.

Values of Δ_{xc} can be obtained from the ensemble method either from the direct calculation of change in the total ensemble energy gradient as the number of electrons in the ensemble system surpasses an integer, or by calculating the difference between the ensemble potentials for $N+\omega$ and $N-\omega$ as $\omega \to 0$. This difference was shown to converge to a constant in the vicinity of the system, decaying to zero as $r \to \infty$. Although the exact Δ_{xc} should be a constant across all space, it is argued that this decay is primarily a result of the Gaussian basis set implementation and not a deficiency of the (semi-)local density functional approximations employed.

Calculated values for the derivative discontinuity via the difference in the ensemble potentials were shown to be too small (although previously they were zero) compared to values determined via the change in gradient of the total ensemble energy. The shift observed in the potentials in this theory is a consequence of the discontinuous change to the screening charge, hence we argue that this accounts for a significant contribution to the total derivative discontinuity.

However, we should note that the value of Q_{scr} does not determine fully Δ_{xc}. It is known that the exact $\Delta_{\mathrm{xc}}(r)$ is a function of r that stays equal to a constant shift in the region of the molecule and eventually drops to zero at a large distance away

from the molecule. The point where it drops to zero is related to the asymptotic tails of the densities of the constituents systems of the ensemble (where they cross and become comparable). We cannot control this effect just with the norm of Q_{scr} and it is also doubtful if (semi-)local DFAs like LDA can describe this, even if we used a grid rather than Gaussian bases.

Chapter 7

Localising the xc screening hole

In this chapter, a novel new method for decomposing the total molecular screening density into screening densities localised on individual atoms is developed, by introducing the concept of atomically localised screening charge. This method builds on the work on previous constrained OEP schemes developed in this thesis, and its ideas could be applied systematically across all of these previous methods.

This work was inspired by a well-known issue with commonly used (semi-)local functionals within DFT, which causes diatomic (or even larger systems) to dissociate incorrectly into fractionally charged species rather than neutral atoms. We will discuss how the derivative discontinuity introduced in § 6, is related to steps in the exact KS potential that form when molecules are stretched; an extremely difficult feature to capture at all using any approximate density functional.

We first discuss results obtained for systems at equilibrium (i.e. small separation distances), to confirm that this new method correctly reduces to the constrained OEP formulation introduced in § 4. Following on from this, we explore results obtained for diatomics that are stretched far beyond their equilibrium bond length, and it is that shifts in the OEP potentials do in fact manifest. Although these shifts are not large enough to completely correct the incorrect fractional charge fragment dissociation of diatomics predicted by (semi-)local density functional approximations, their observation alone is nevertheless a spectacular result.

7.1 Theory

7.1.1 The fractional dissociation error for (semi-)local functionals

In the previous chapter, we discussed how commonly used (semi-)local density functional approximations do not possess an xc derivative discontinuity as the number of electrons in a system surpasses an integer number of electrons. This deficiency is linked to self-interaction errors that plague these functionals, which causes the system's electronic density to over delocalise. As a result, this error is commonly known as the delocalisation error [131, 132], and it is responsible for other issues with (semi-)local functionals; notably, the tendency for diatomic molecules to energetically favour dissociating to fractionally charged fragments [104].

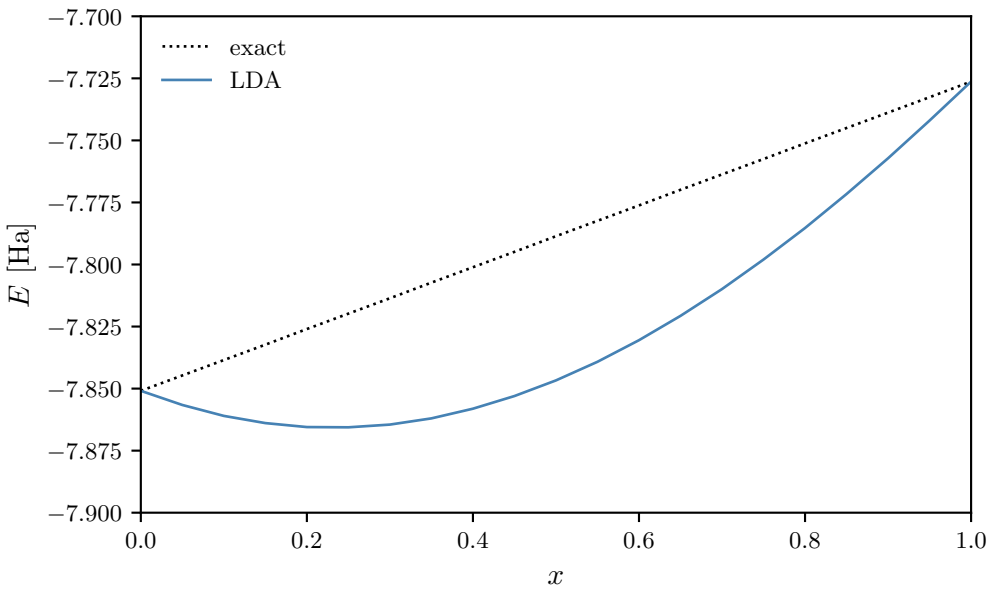


Figure 7.1: The energy of a dissociated LiH molecule calculated using the LDA functional, obtained from the sum of individual calculations of Li^{+x} and H^{-x}, $0 \leq x \leq 1$. The dotted line represents the expected linear result that would be calculated from the exact KS potential. Each of these calculations was performed using an augmented cc-pVTZ orbital basis set.

The important work of Perdew et al. [104] highlights this error by studying the dissociation of the LiH molecule, where the LDA displays no derivative discon-

tinuity or steps in the xc potential and therefore incorrectly minimises the total ground state energy at $Li^{+0.25}H^{-0.25}$, rather than the correct dissociation limit of two neutral atoms of Li and H (see Figure 7.1). This is caused by the fact that in (semi-)local DFAs, the electron affinity (EA) of one atom, here H, is considerably overestimated compared to the ionisation potential (IP) of the other atom, here Li. It is not just approximations in DFT that suffer from fractional dissociation errors, however, as restricted HF in fact yields an even worse limit of $Li^{+0.45}H^{-0.45}$ [104].

Clearly, the exact KS potential should not exhibit this behaviour, and should instead dissociate diatomics to neutral atoms. To achieve this, the exact KS potential develops step-like features which prevent the energy minimum from yielding charged fragments [133]. These steps in the KS potential have also recently been related to the xc derivative discontinuity by Hodgson et al. [134].

Accurately capturing these steps is of vital importance when studying systems where the atoms (or separate parts of a molecule) appear as individual entities, such as a molecule with a donor and an acceptor site, or a molecule near the surface of a solid.

Unfortunately, the ability to capture these steps using (semi-)local density functional approximations so far has been impossible to tackle. Therefore, we must somehow develop alternative approaches that can incorporate these steps into existing DFAs.

7.1.2 OEP with atomically localised screening densities

In the previous implementations of the OEP equation discussed in this thesis, the effective potential is expanded in terms of a screening density distributed across the entire space of the molecule. Instead, we introduce here the concept of an atomic screening density localised on a particular atom in the molecule, where the total screening density of the molecule is given by the sum of all the individual atomic

screening densities:

$$\rho_{\text{scr}}^{\text{mol}}(\mathbf{r}) = \sum_{I=1}^{M} \rho_{\text{scr}}^{I}(\mathbf{r}), \tag{7.1}$$

where M is the number of atoms in the molecule and $\rho_{\text{scr}}^{I}(\mathbf{r})$ is the atomic screening density of atom I.

As before, the effective potential can be expanded in terms of the total molecular screening density, and using the above expression it can be further split up into individual atomic effective potentials each expanded in terms of their respective atomic screening density:

$$v_{\text{eff}}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \frac{\rho_{\text{scr}}^{\text{mol}}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \sum_{I=1}^{M} \int \mathrm{d}\mathbf{r}' \frac{\rho_{\text{scr}}^{I}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \sum_{I=1}^{M} v_{\text{eff}}^{I}, \tag{7.2}$$

where:

$$v_{\text{eff}}^{I}(\mathbf{r}) = \int \mathrm{d}\mathbf{r}' \frac{\rho_{\text{scr}}^{I}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{7.3}$$

Following the ideas developed in § 4, we know that the molecular Hxc screening density should be constrained to integrate to $N - 1$ to yield the correct asymptotic decay of the OEP potential at infinity and to correct self-interaction errors that contaminate commonly used (semi-)local density functional approximations. This, therefore, allows us to constrain the sum of the individual atomic screening densities, such that:

$$\int \mathrm{d}\mathbf{r}' \rho_{\text{scr}}^{\text{mol}}(\mathbf{r}') = \sum_{I=1}^{M} \int \mathrm{d}\mathbf{r}' \rho_{\text{scr}}^{I}(\mathbf{r}') = \sum_{I=1}^{M} Q_{\text{scr}}^{I} = N - 1. \tag{7.4}$$

If we consider what happens to a diatomic when it is stretched, it is expected that the xc screening hole of charge -1 should localised onto one of the atoms (physically, this should be the atom with the most loosely bound or diffuse electron). Therefore, only one of the atomic screening charges, $Q_{\text{scr}}^{J} = N^{J} - 1$, with all of the other $Q_{\text{scr}}^{I \neq J} = N^{I}$.

This argument regarding the localisation of the -1 hole strictly holds at dissociation but here we employ it in general. This is because at infinity the density of that particular electron dominates and it should see $N - 1$ electrons of its own atom to repel it (SI-free). So, we choose to localise the -1 hole on the atom with the

smallest IP. For a homonuclear diatomic, the -1 hole should still localise on one of the atoms; the one which is closest to the point r we are considering at infinity.

To continue, each of the atomic screening densities are expanded on the set of auxiliary basis functions $\{\theta_n^I\}$, that are localised to that particular atom:

$$\rho_{\text{scr}}^I(\mathbf{r}) = \sum_{n=1}^{N^I} c_n^I \theta_n^I(\mathbf{r}), \tag{7.5}$$

where the set of basis functions localised to atom I, $\{\theta_1^I, \ldots, \theta_{N^I}^I\}$, is a subset of the total set of basis functions for the whole molecule.

In Figure 7.2, it is clear that at equilibrium distances the basis set functions corresponding to each atom in the LiH molecule have significant overlap; however, at large separations, this overlap decreases dramatically. It should therefore be expected (and by design) that the localisation of screening charges to individual atoms should have a maximal impact when the atoms are infinitely separated, i.e. dissociated.

The constraints on atomic screening charges in terms of the atomic screening density expansion coefficients are therefore:

$$\sum_{n=1}^{N^I} c_n^I \int \mathrm{d}\mathbf{r}\, \theta_n^I(\mathbf{r}) = \sum_{n=1}^{N^I} c_n^I \, X_n^I = Q_{\text{scr}}^I, \tag{7.6}$$

where:

$$X_n^I = \int \mathrm{d}\mathbf{r}\, \theta_n^I(\mathbf{r}), \quad n = 1, \ldots, N^I. \tag{7.7}$$

It should be clear that the total number of coefficients over all the atoms will equal the total number of basis functions for the molecule; the set coefficients $\{c_n^I\}$ are simply a subset of the total set of coefficients that describe the screening density localised to atom I. By varying one of these coefficients, c_n^I, we introduce a change in the atomic screening density for the atom I:

$$\delta\rho_{\text{scr}}^I(\mathbf{r}) = \delta c_n^I \theta_n^I(\mathbf{r}), \tag{7.8}$$

which in turn results in an overall change in the total energy functional, that we will minimise under the set of constraints $-\sum_J^M \lambda^J \int \mathrm{d}\mathbf{r} \rho_{\text{scr}}^J(\mathbf{r})$. The change in total

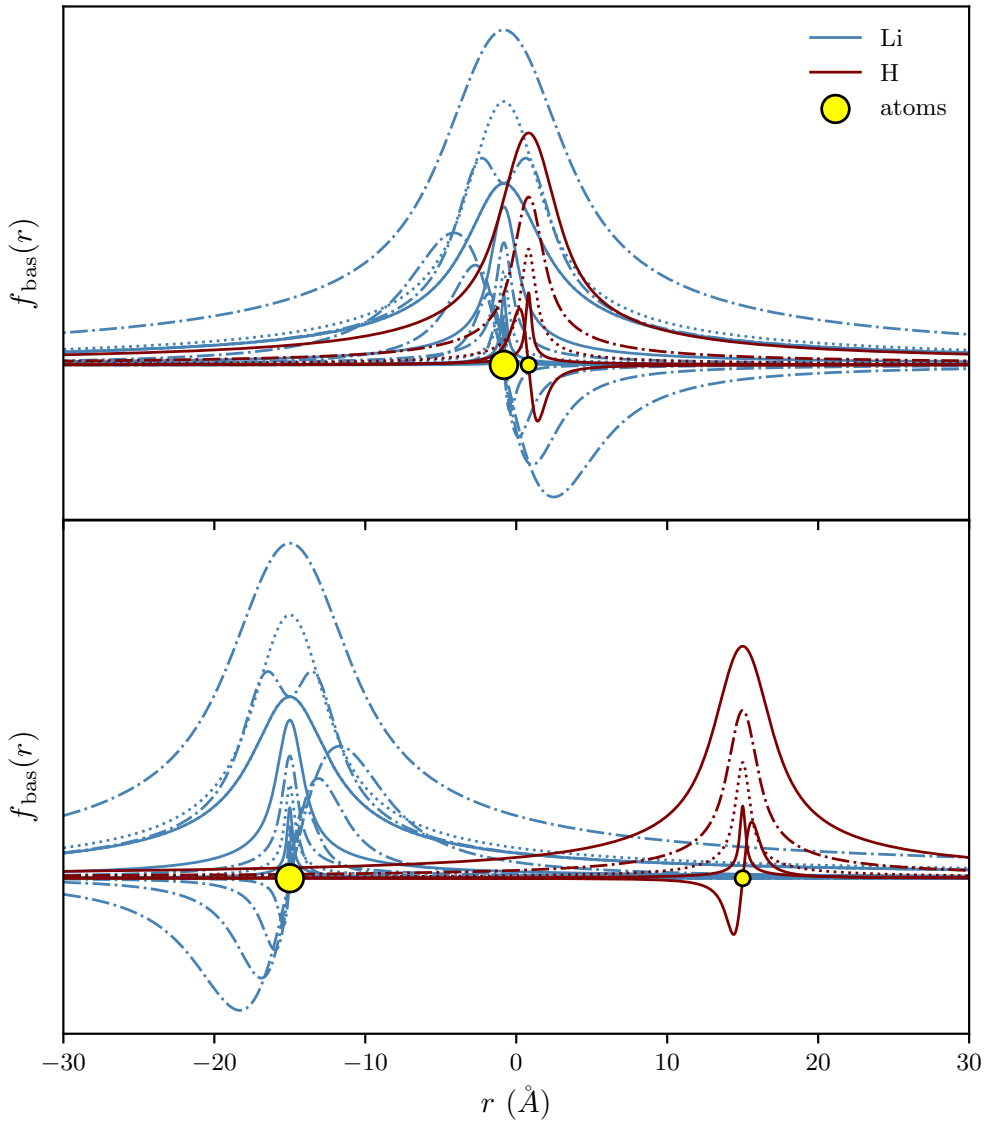


Figure 7.2: The cc-pVDZ basis set functions for a molecule of LiH at two separation distances; (top) equilibrium distance and (bottom) 30Å. The two colours represent which functions are localised to each atom; with the Lithium functions coloured blue, and the Hydrogen functions in red. The two yellow markers indicate the relative positions of the two atoms.

energy functional under these constraints is given by:

$$\Delta E - \lambda^I \int \mathrm{d}\mathbf{x}' \delta\rho^I_{\mathrm{scr}}(\mathbf{x}') = 0 \tag{7.9}$$

$$\implies \iiint \mathrm{d}\mathbf{x}'\mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{r}\ \frac{\delta\rho^I_{\mathrm{scr}}(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\chi(\mathbf{r},\mathbf{x})\Bigg[v^{\mathrm{DFA}}_{\mathrm{Hxc}}(\mathbf{r}) - \sum_{J=1}^{M} v^J_{\mathrm{eff}}(\mathbf{r})\Bigg] - \lambda^I \int \mathrm{d}\mathbf{x}'\delta\rho^I_{\mathrm{scr}}(\mathbf{x}') = 0. \quad (7.10)$$

Performing the functional derivative on this above expression yields a familiar-looking equation:

$$0 = \int \mathrm{d}\mathbf{x}'\theta^I_n(\mathbf{x}') \iint \mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{r}\ \frac{\chi(\mathbf{r},\mathbf{x})}{|\mathbf{x}-\mathbf{x}'|}\Big[v^{\mathrm{DFA}}_{\mathrm{Hxc}}(\mathbf{r}) - \sum_{J=1}^{M} v^J_{\mathrm{eff}}(\mathbf{r})\Big] - \lambda^I \int \mathrm{d}\mathbf{x}'\theta^I_n(\mathbf{x}'). \quad (7.11)$$

The above equation can be rewritten as:

$$\iint \mathrm{d}\mathbf{x}'\mathrm{d}\mathbf{r}'\ \theta^I_n(\mathbf{x}')\bar{\chi}(\mathbf{x}',\mathbf{r}') \sum_{J=1}^{M}\sum_{l=1}^{N^J} c^J_l\theta^J_l(\mathbf{r}') = b^I_n - \lambda^I X^I_n, \quad (7.12)$$

where for $n = 1,\ldots,N^I$:

$$b^I_n = \iiint \frac{\mathrm{d}\mathbf{x}'\mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{r}}{|\mathbf{x}-\mathbf{x}'|}\chi(\mathbf{r},\mathbf{x})v^{\mathrm{DFA}}_{\mathrm{Hxc}}(\mathbf{r})\theta^I_n(\mathbf{x}'). \quad (7.13)$$

By defining the matrix $A_{In;Jl} = \left\langle \theta^I_n \right| \bar{\chi} \left| \theta^J_l \right\rangle$, the OEP equation can once again be re-expressed as in a matrix form:

$$\sum_{J=1}^{M}\sum_{l=1}^{N^J} A_{In;Jl}\, c^J_l = b^I_n - \lambda^I\, X^I_n; \quad n = 1,\ldots,N^I. \quad (7.14)$$

Multiplying by the inverse of A (same matrix as before) and summing over I and $n = 1,\ldots,N^I$ yields:

$$c^K_m = \sum_{I=1}^{M}\sum_{n=1}^{N^I} A^{-1}_{Km;In}b^I_n - \sum_{I=1}^{M}\lambda^I \sum_{n=1}^{N^I} A^{-1}_{Km;In}X^I_n; \quad m = 1,\ldots,N^K. \quad (7.15)$$

To complete this reformulation of the OEP, we must construct expressions for the Lagrange multipliers, λ_J. To do so is straightforward, as multiplying Equation 7.15 by X^K_m with $m = 1,\ldots,N^K$, and summing over m yields:

$$\sum_{m=1}^{N^K} c^K_m X^K_m = \sum_{I=1}^{M}\sum_{n=1}^{N^I}\sum_{m=1}^{N^K} X^K_m\, A^{-1}_{Km;In}b^I_n - \sum_{I=1}^{M}\lambda^I \sum_{n=1}^{N^I}\sum_{m=1}^{N^K} X^K_m A^{-1}_{Km;In}X^I_n \quad (7.16)$$

$$(7.17)$$

$$\implies Q^K_{\mathrm{scr}} = B^K - \sum_{I=1}^{M}\tilde{A}_{KI}\,\lambda^I, \quad (7.18)$$

where the matrices $\tilde{A}_{KI}$ and B_K are given by:

$$\tilde{A}_{KI} = \sum_{n=1}^{N^I} \sum_{m=1}^{N^K} X_m^K A_{Km;In}^{-1} X_n^I, \tag{7.19}$$

$$B^K = \sum_{I=1}^{M} \sum_{n=1}^{N^I} \sum_{m=1}^{N^K} X_m^K \, A_{Km;In}^{-1} b_n^I. \tag{7.20}$$

The Lagrange multipliers can then be determined by inverting this equation:

$$\lambda^J = \sum_{K=1}^{M} \tilde{A}_{JK}^{-1} (B^K - Q_{\mathrm{scr}}^K). \tag{7.21}$$

This novel method for localising the screening density in the OEP potential has been implemented into the HIPPO code in combination with the implicit OEP formulation. It can be applied in principle to any (semi-)local density functional approximation, however, for this work, only the LDA and PBE functionals were studied. The code has the flexibility to localise the xc screening density hole of -1 on any atom of choice. This allows us to perform several calculations with the xc screening hole confined to different atoms.

7.1.3 Non-overlapping atomic screening densities

One of the limitations of the atomically localised OEP method presented in the previous section is that in the limit that the auxiliary basis centered on each atom becomes very large (i.e. towards complete), the basis functions of each atom will always overlap with those of every other atom. This, therefore, leads to an ambiguity as to where to assign the screening charges on atoms, and in this limit, it reduces to the constrained or implicit OEP that models the screening density as being smeared out across the entire molecule. In what follows, an additional constraint on the atomic screening densities is introduced and the atomically localised OEP equation is modified to incorporate it. However, for the size of the basis sets used both in the work of this thesis and commonly used by the electronic structure community, this constraint is not strictly necessary; nevertheless, it is important to address this limit.

The additional condition that should be enforced on the atomic screening densities to avoid this issue is that none of the atomic screening densities should overlap with each other:

$$\prod_{J=1}^{M} \rho_{\mathrm{scr}}^{J}(\mathbf{r}) = 0; \quad \forall\ \mathbf{r}, \tag{7.22}$$

or equivalently:

$$\int \mathrm{d}\mathbf{r} \prod_{J=1}^{M} |\rho_{\mathrm{scr}}^{J}(\mathbf{r})| = 0. \tag{7.23}$$

Like the positivity constraint introduced in this thesis, this non-overlapping condition can be incorporated into the objective function as a penalty term:

$$+\ \Lambda \int \mathrm{d}\mathbf{r} \prod_{J=1}^{M} |\rho_{\mathrm{scr}}^{J}(\mathbf{r})|; \quad \Lambda \in \mathbb{R}^{+}. \tag{7.24}$$

The change in the total energy functional from Equation 7.9 including this additional term is therefore now:

$$\begin{aligned} &\Delta E - \lambda^{I} \int \mathrm{d}\mathbf{x}' \delta\rho_{\mathrm{scr}}^{I}(\mathbf{x}') \\ &\qquad + \Lambda \int \mathrm{d}\mathbf{x}' \delta\rho_{\mathrm{scr}}^{I}(\mathbf{x}')\ \mathrm{sgn}\big[\rho_{\mathrm{scr}}^{I}(\mathbf{x}')\big] \prod_{J\neq I}^{M} |\rho_{\mathrm{scr}}^{J}(\mathbf{x}')| = 0, \end{aligned} \tag{7.25}$$

where $\mathrm{sgn}\big[\rho_{\mathrm{scr}}^{I}(\mathbf{x}')\big]$ is the sign of $\rho_{\mathrm{scr}}^{I}(\mathbf{x}')$ at $\mathbf{x}'$.

Performing the functional derivative then yields:

$$\begin{aligned} 0 \;=\; &\int \mathrm{d}\mathbf{x}' \theta_{n}^{I}(\mathbf{x}') \iint \mathrm{d}\mathbf{x}\ \mathrm{d}\mathbf{r}\ \frac{\chi(\mathbf{r},\mathbf{x})}{|\mathbf{x}-\mathbf{x}'|} \Big[v_{\mathrm{Hxc}}^{\mathrm{DFA}}(\mathbf{r}) - \sum_{J}^{M} v_{\mathrm{eff}}^{J}(\mathbf{r}) \Big] \\ &- \lambda^{I} \int \mathrm{d}\mathbf{x}' \theta_{n}^{I}(\mathbf{x}') + \Lambda \int \mathrm{d}\mathbf{x}' \theta_{n}^{I}(\mathbf{x}')\, \mathrm{sgn}\big[\rho_{\mathrm{scr}}^{I}(\mathbf{x}')\big] \prod_{J\neq I}^{M} |\rho_{\mathrm{scr}}^{J}(\mathbf{x}')|, \end{aligned} \tag{7.26}$$

which after some rearrangement can be rewritten in the regular OEP form as:

$$\iint \mathrm{d}\mathbf{x}'\, \mathrm{d}\mathbf{r}'\ \theta_{n}^{I}(\mathbf{x}') \bar{\chi}(\mathbf{x}',\mathbf{r}') \sum_{J=1}^{M} \sum_{l=1}^{N^{J}} c_{l}^{J} \theta_{l}^{J}(\mathbf{r}') = b_{n}^{I} - \lambda^{I} X_{n}^{I} + \Lambda \bar{X}_{n}^{I}, \tag{7.27}$$

for $n = 1, \ldots, N^{I}$, b_{n}^{I} given by the same form as in Equation 7.13, and where:

$$\bar{X}_{n}^{I} = \int \mathrm{d}\mathbf{r}\, \theta_{n}^{I}(\mathbf{r})\ \mathrm{sgn}\big[\rho_{\mathrm{eff}}^{I}(\mathbf{r})\big] \prod_{J\neq I}^{M} |\rho_{\mathrm{scr}}^{J}(\mathbf{r})|. \tag{7.28}$$

The matrix form of this equation is therefore given by:

$$\sum_{J}^{M} \sum_{l=1}^{N^{J}} A_{In;Jl}\, c_{l}^{J} = b_{n}^{I} - \lambda^{I} X_{n}^{I} + \Lambda \bar{X}_{n}^{I}, \tag{7.29}$$

The above equation now provides a version of the atomically localised screening density OEP equation that allows the screening densities to be localised optimally onto individual atoms even in the complete basis set limit. As stated before however, for the systems studied and basis sets used for the work in this thesis, the inclusion of this constraint did not have an impact on results; it is however important to present a method of overcoming this limitation that may prove useful in the future.

7.2 Results

7.2.1 Equilibrium geometries

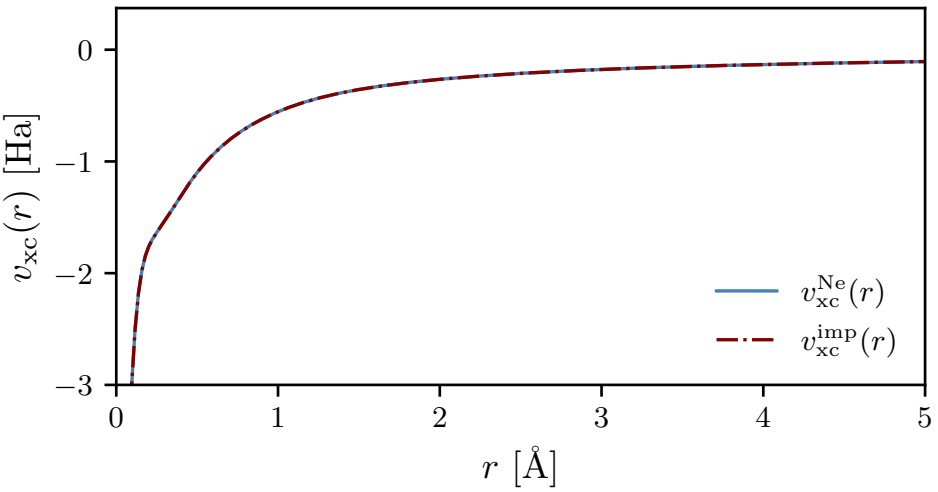


Figure 7.3: The exchange-correlation potentials for Neon (Ne) and calculated using implicit and the atomically localised xc screening hole method for the LDA functional. Both calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

In this section, we will explore the behaviour of this new OEP reformulation for systems with small atomic separations. As was mentioned earlier, for such cases the separation of the total screening density into separation densities localised to individual atoms should reduce to the implicit OEP scheme, as there will be a significant overlap between the subsets of the auxiliary basis functions localised on

each atom with all of the other subsets. Certainly, in the case of a single atom, results should be identical, as then the atomically localised auxiliary basis functions includes the entire set.

In Figure 7.3 we perform this sanity check on the Neon (Ne) atom, by plotting the xc potentials obtained from both implicit LDA and the new atomically localised xc screening hole scheme for the LDA functional. It is clear from the plot that the two potentials are identical (within numerical noise); a result that was expected and is now confirmed.

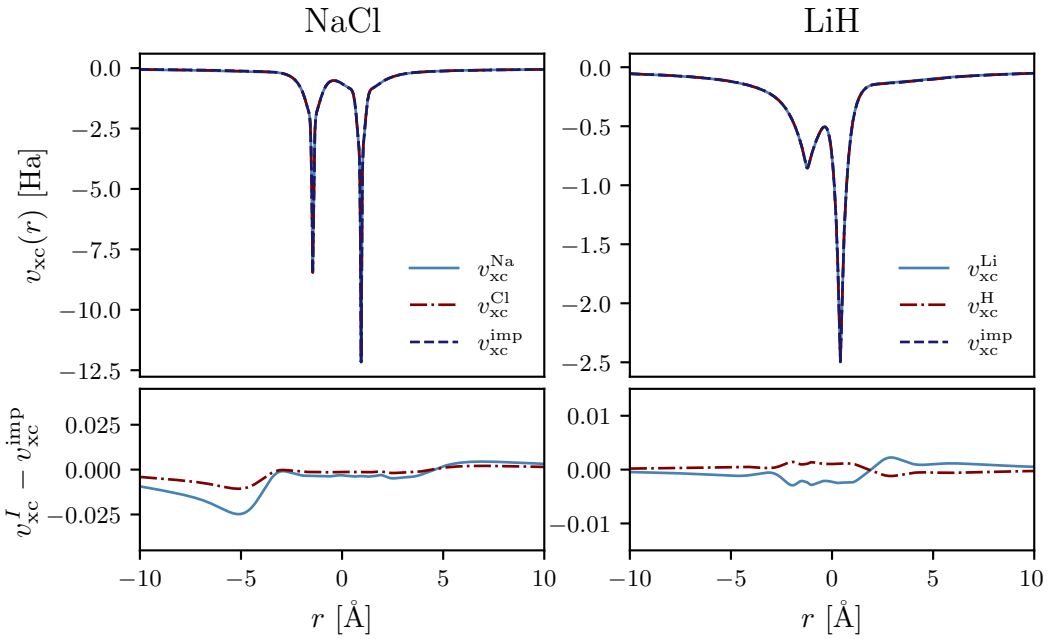


Figure 7.4: The exchange-correlation potentials for NaCl and LiH were calculated using the implicit and atomic OEP methods. The bottom subplots show the difference between the atomic OEP xc potential with charge localised on atom I and the implicit potential. All calculations used an augmented cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

More interestingly, we can study the behaviour of the two methods for some diatomic systems at their equilibrium configurations. In Figure 7.4, we again compare the implicit LDA xc potential to the new method however this time we have two possibilities of where to localise the xc screening hole in each system (although strictly it should only localise on the atom with the smallest IP). In the left-hand

side of the figure, the three xc potentials obtained for NaCl are plotted, along with the differences between the potentials that are calculated with the xc screening hole localised on the Na and Cl atoms respectively, and the potential calculated using the implicit method. On the right-hand side of the figure, the same potentials and differences and plotted but instead for LiH.

From the two bottom subplots, it is clear that the difference between the potentials obtained with the xc screening hole localised on one of the atoms and the implicit LDA potential is small (at most ~ 0.02 Ha for NaCl and ~ 0.002 Ha for LiH).

For LiH, there is not a significant difference between the potentials when the hole is localised on either of the two atoms. However, for NaCl we do observe minor differences on the Na side of the potential when the xc screening hole is localised to the Na atom (the atom with the smallest affinity of 0.547 eV [135]). As it is energetically favourable, the hole naturally wants to localise strongly to the Cl atom (which has a much higher affinity of 3.612 eV [136]), the potential that results from the Cl localisation is very similar to the implicit LDA result. Therefore, even at the equilibrium distance, we can start to see a difference between the potentials depending on where the hole is localised, and it should be localised on the atom with the smallest IP (in this case Na). This could be explained by the relative differences between the affinities in NaCl of 0.136 eV compared to the difference between the H affinity (0.754 eV [137]) and the Li affinity (0.618 eV [138]). Although, it should be noted that for both systems the total energy differences between these solutions is in fact very small ($\sim 10^{-6}$ Ha).

Together these figures highlight the expected behaviour of this approach; that for small distances the localisation of the xc screening charge hole onto an individual atom has only a small effect on the resulting potentials as there is significant overlap of the atomically localised sets of auxiliary basis functions. Although, for diatomics with larger differences in the electron affinities of the two atoms minor steps begin to develop.

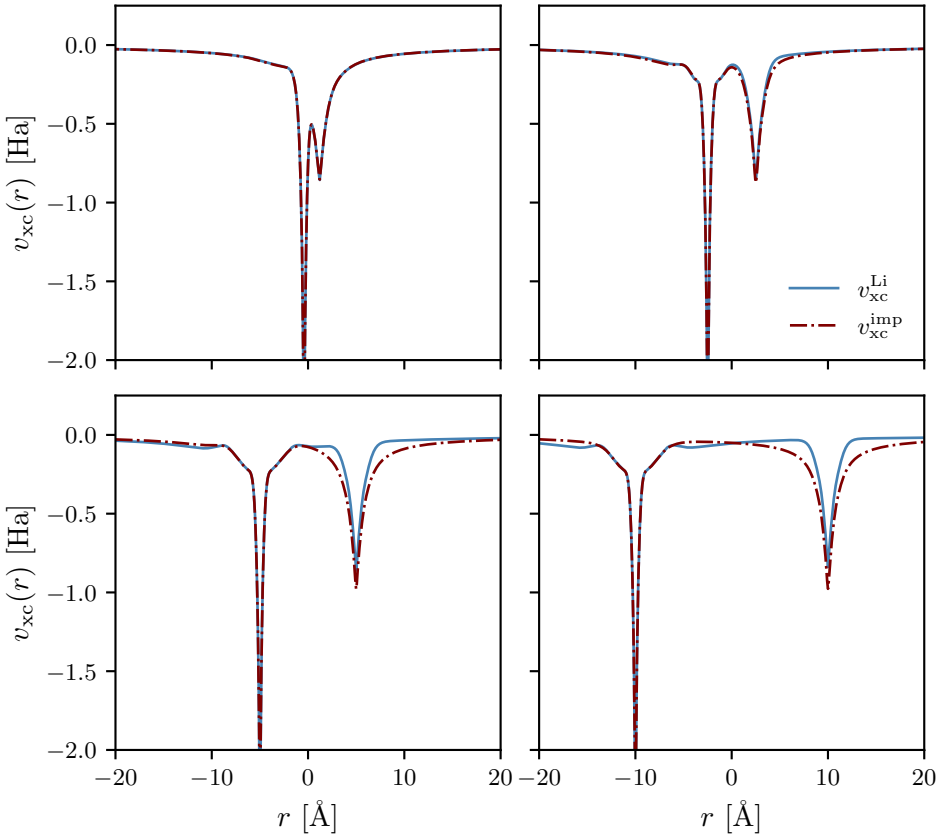


Figure 7.5: The exchange-correlation potentials for LiH calculated via implicit LDA and for the xc screening hole localised on the Li atom for four separation distances, d: (top left) equilibrium distance, $d = 1.64$Å, (top right) $d = 5.0$Å, (bottom left) $d = 10.0$Å, and (bottom right) $d = 20.0$Å. All calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

7.2.2 The emergence of steps in the xc-potential

We have so far confirmed that decomposing the molecular screening charge into screening charges localised on individual atoms reduces to the regular constrained OEP method at small diatomic separation distances. We can now test the method as molecules are stretched far beyond their equilibrium bonding length to see whether or not the potentials calculated show any systematic variations.

In Figure 7.5, we plot the OEP potentials for LiH obtained via implicit OEP

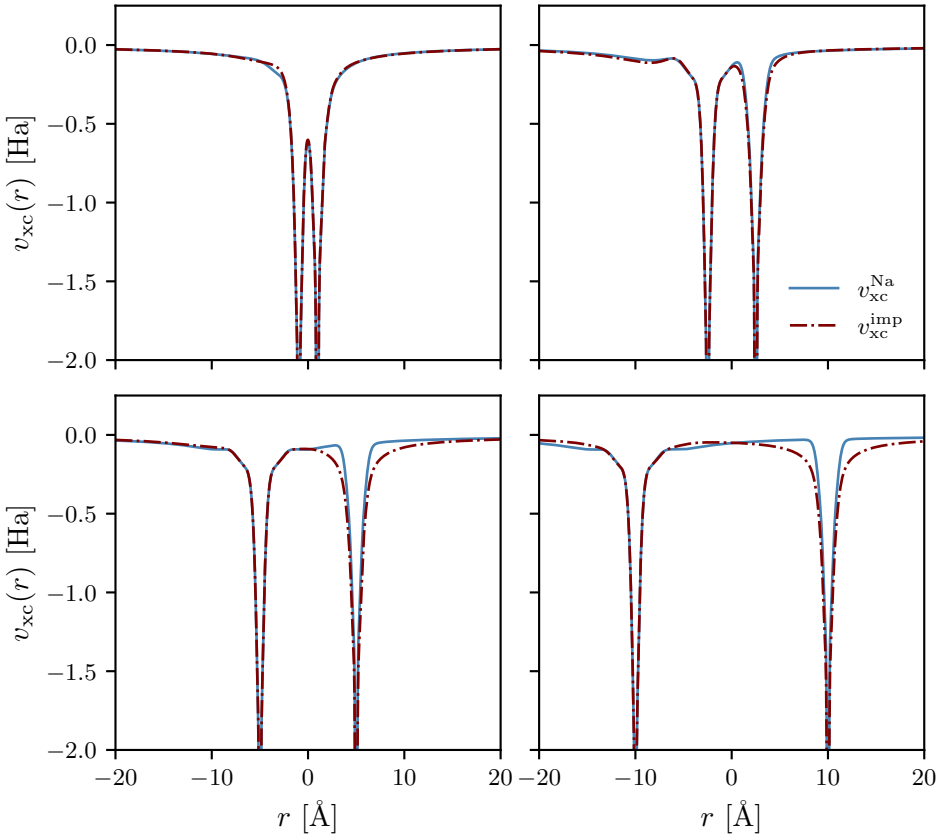


Figure 7.6: The exchange-correlation potentials for NaF calculated via implicit LDA and for the xc screening hole localised on the Na atom for four separation distances, d: (top left) equilibrium distance, $d = 1.93$Å, (top right) $d = 5.0$Å, (bottom left) $d = 10.0$Å, and (bottom right) $d = 20.0$Å. All calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

and the new method with the xc screening charge hole localised on the Lithium atom (the smallest IP), for four different separation distances, ranging from the equilibrium distance of 1.64Å, up to 20.0Å (the potential obtained with the xc screening hole localised on the Hydrogen atom does not show significant deviations from the implicit potential). It is clear from these figures, that as the separation distance increases, there is a significant change between the potentials determined from both methods, with the new method appearing to shift the potential in the region of the Hydrogen atom upwards.

d [Å]	Q(Li)	Q(H)
1.64	2.6714	1.3286
5.00	3.0015	0.9985
10.00	2.9999	1.0001
20.00	3.0000	1.0000

d [Å]	Q(Na)	Q(F)
1.93	10.1412	9.8588
5.00	10.9993	9.0007
10.00	10.9999	9.0001
20.00	11.0000	9.0000

Table 7.1: Calculated Mulliken charges, $Q(I)$, for increasing separation distance, d, for the atoms of (left) LiH with the xc screening hole localised on the Li atom and (right) NaF with the xc screening hole localised on the Na atom. All calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

A very similar picture emerges in Figure 7.6, where we again plot the xc potentials obtained from the implicit OEP and the new atomically localised OEP method for four increasingly large separation distances, this time for NaF, with the xc screening charge hole localised to the Na atom (the smallest IP). Here, the shift in the xc potential as the separation distance between the two atoms increases also manifests; this time in the vicinity of the F atom.

Using Mulliken population analysis [139], we were able to confirm these self-consistent solutions to the OEP equation for both systems yield neutral atoms at large separation distances (see Table 7.1). However, without a strict energy minimisation routine, we cannot be completely sure that these solutions are the lowest possible ground state energy solutions; this requires a more detailed investigation in the future.

These shifts are expected because at dissociation and far away from both atoms, the 3s electron of Na sees 10/11 electrons of Na to repel it (i.e. no SI) as well as all 9/9 electrons of F. Therefore, the potential of F is shifted upwards in the eyes of the 3s electron of Na, compared to the corresponding potential seen by a 2p electron of F at infinity.

The inter-atomic steps are a feature of the exact KS potential which prevents these diatomics from dissociating to unphysical fractionally charged fragments, and they are extremely difficult to capture at all with (semi-)local density functional approx-

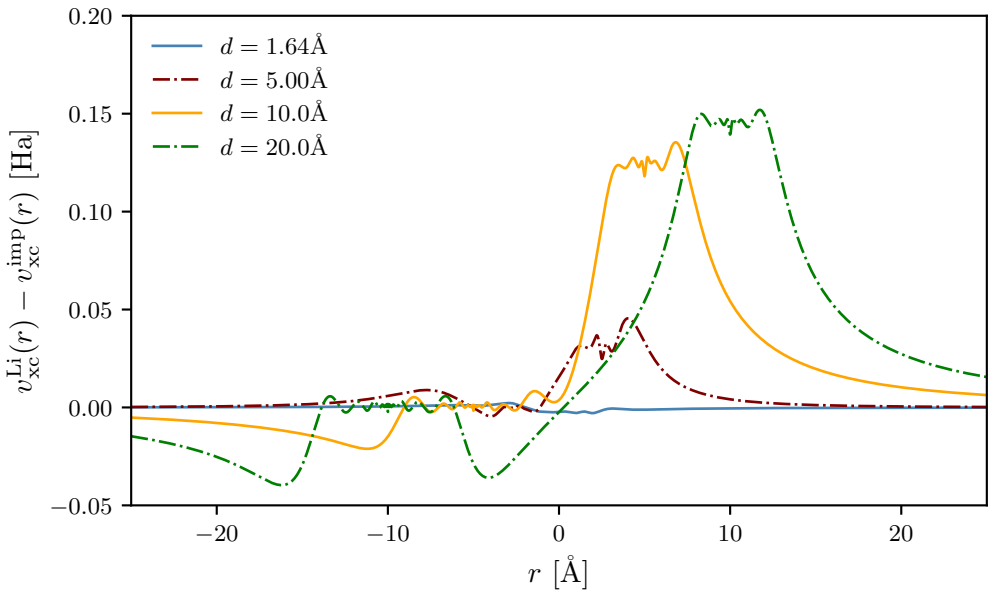


Figure 7.7: Differences between the OEP xc potential when the xc screening charge hole is localised on the Li atom using the LDA functional and the implicit LDA result for LiH, for four separation distances (with the Li atom located at $-d/2$ and the H atom at $+d/2$). All calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

imations. As these results were obtained using simply the LDA approximation, the presence of any kind of shift is a fairly remarkable result, as in previous implementations no such steps have ever been observed.

We can study the steps for these two examples in more detail, by plotting the differences between the implicit OEP potentials and those calculated with the xc screening hole localised on the Li and Na atoms respectively.

Figure 7.7 plots the differences in the two potentials for LiH for each of the four separation distances presented in 7.5. At equilibrium, we notice that the two potentials are nearly identical, with only very minor derivations between the two observed; in agreement with results obtained in the previous section. As the separation distance increases to 5Å, a small shift in the potentials in the vicinity of the Hydrogen atom appears. At this point, the overlap between the two sets of auxiliary basis functions localised on each atom has decreased sufficiently for a

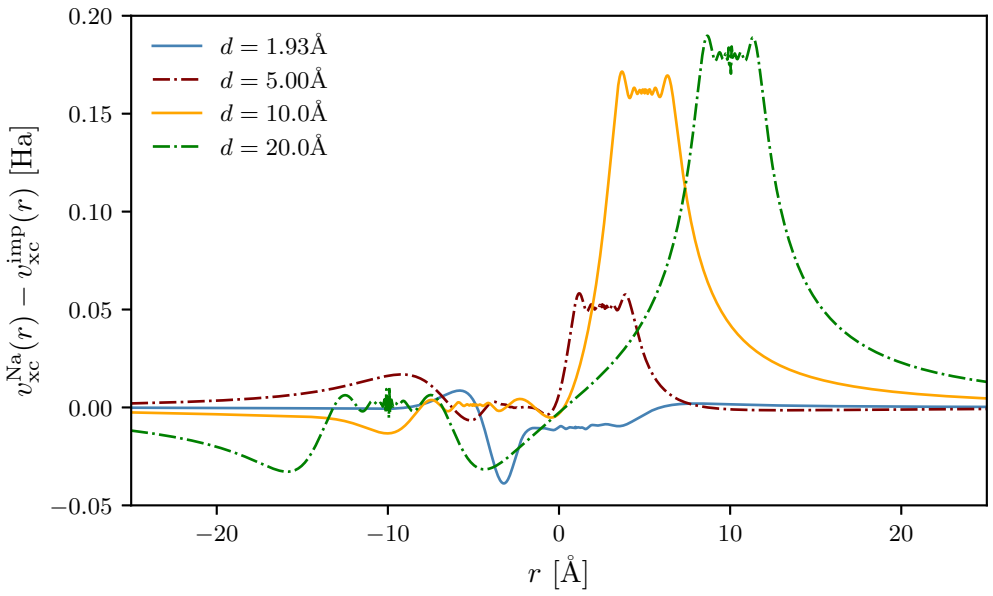


Figure 7.8: Differences between the OEP xc potential when the xc screening charge hole is localised on the Na atom using the LDA functional and the implicit LDA result for NaF, for four separation distances (with the Na atom located at $-d/2$ and the F atom at $+d/2$). All calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

different self-consistent solution to the OEP equation to be obtained. When the separation distance reaches 10Å, this shift begins to appear more as a flat step in the neighbourhood of the Hydrogen atom (similar to the step observed in § 6.2.2), and at 20Å this step plateaus at approximately 0.15 Ha.

In Figure 7.8 we plot the differences for NaF in the same manner. Once again, at equilibrium only minor variations between the two potentials are observed; these are however more pronounced than for LiH, likely due to the relatively larger differences between the affinities for Na and F. As the diatomic is stretched further, a flat step-like feature begins to manifest in the vicinity of the Fluorine atom.

In both cases, the step is local to the atoms with the larger affinity and where the xc screening hole is not localised. These steps essentially act as a penalty for any extra charge which tries to localise onto them, in an attempt to prevent the dissociation into fractionally charged species from becoming the energetically

favourable solution. In the following section, we will discuss whether the magnitude of these observed steps is sufficient to correct this error, however here we note again that simply the presence of these steps regardless of their height is significant. These steps are predicted feature of the exact KS potential but have previously never been observed in the previous finite basis implementations of the OEP scheme for (semi-)local functionals developed in this thesis.

7.2.3 Correcting the fractional charge fragment dissociation error

In this section, we will discuss whether methods developed in this thesis are sufficient to correct the fractional dissociation error that plagues (semi-)local functionals.

In Figure 7.9, we plot the total energies obtained from ensemble LDA, implicit LDA, and regular LDA, for isolated atoms with a varying number of electrons. In the first row, we plot the energies for the atoms with $N - x$ electrons, and in the second row we plot energies for the atoms with $N + x$ electrons. In the third row, we then sum the energies of the two corresponding above subplots, to calculate the energy of the respective dissociated diatomic, in an attempt to reproduce and understand the results obtained by Perdew et al. [104] (although this paper only studies LiH in this manner).

For each of the systems presented in these subplots, the ensemble OEP yields a linearly varying ground state total energy and correctly predicts that the lowest energy state for the dissociation of all three molecules is the case of neutral atoms. When we instead look at results obtained via iLDA and LDA, we notice the convex shape of the energy curves that were discussed in § 6.1.2, and the incorrectly predicted dissociation into fractionally charged species.

Focusing on the bottom-left subplot, we can see however that in the energy curves calculated for the dissociation of LiH, iLDA does in fact yield a minimum at $x = 0$, i.e. neutral atoms. The difference here is due to the ghost exchange error, which

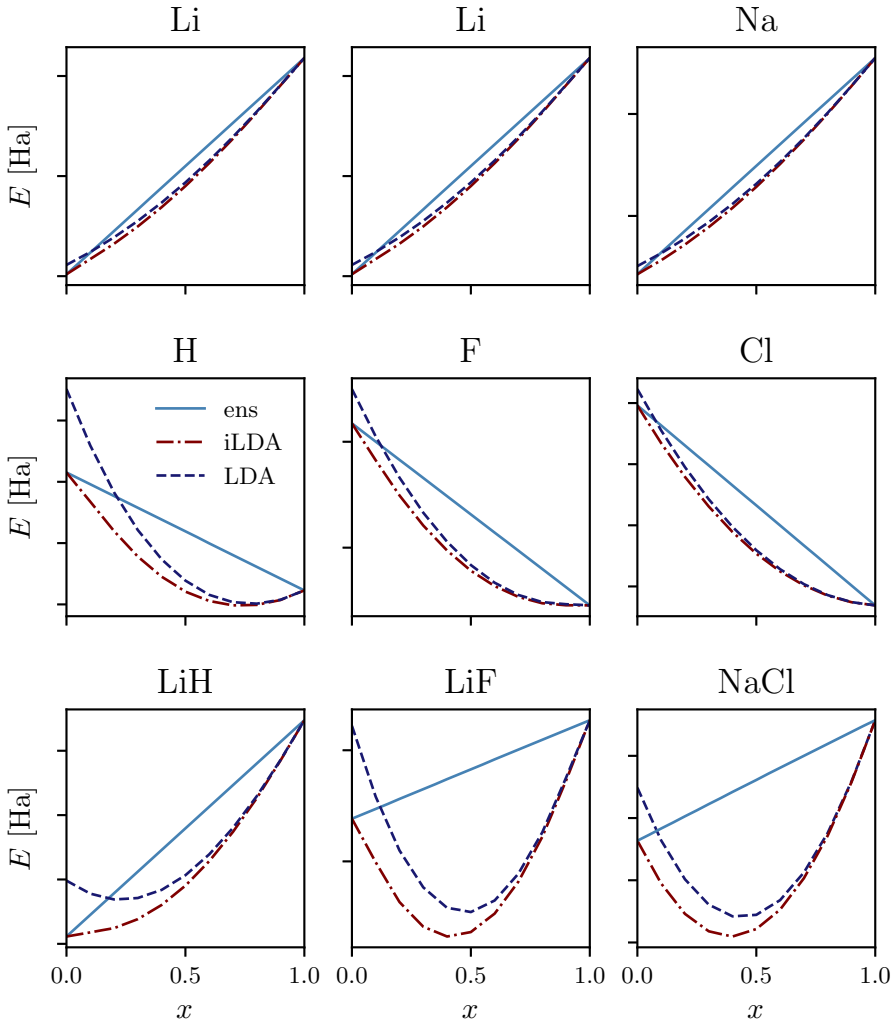


Figure 7.9: Total ground state energies obtained from ensemble LDA, implicit LDA, and regular LDA for systems with a variable number of electrons. In the first row, the energies are calculated for the $N - x$ atomic system, whereas in the second row we plot total energies for the atom with $N + x$ electrons. In the third row, we plot the sum of the total energies from the two above subplots in each column, yielding the total energy for the dissociated diatomic. From this row we can then determine the value of x which gives the minimum ground state total energy for the diatomic at infinite separation distance. All calculations were performed using a cc-pVTZ orbital basis set and an uncontracted cc-pVTZ auxiliary basis set.

for this system is sufficient to completely correct this unphysical dissociation. For LiF and NaCl, the ghost exchange error is clearly not large enough to completely correct for this; although, the value of x that yields the lowest energy is shifted closer to zero.

The localisation of the screening charge hole on the atom with the smallest IP was introduced to model the fact that at dissociation the most loosely bound electron will be repelled by $N-1$ electrons of its own atom. This leads naturally to an upward shift of the xc potential of the other atom, introducing an inter-atomic step. However, even with this step, the EA of F and Cl are still overestimated (see Figure 7.9), and unfortunately the EAs of F and Cl are still larger than the IPs of Li and Na. Nevertheless, we do get closer to the true physical behaviour.

We can therefore conclude that the ghost exchange error alone is not enough to completely correct the dissociation limits for every system; although it can somewhat decrease the amount of spurious fractional charge on the resulting dissociated species.

7.3 Summary and discussion

In this chapter, a novel new method for localising the xc screening hole was formulated, by introducing the concept of atomic screening densities localised on individual atoms. This idea led to the derivation of a modified OEP scheme, which was then implemented into the HIPPO finite basis set code.

The results obtained from this method are very promising. At small separation distances (and for calculations on single atoms), potentials obtained from the atomically localised OEP were equal to those obtained using the constrained or implicit OEP methods for closed-shell systems. This observation is to be expected, as at small separations the auxiliary basis set functions for all of the atoms in the molecule have significant overlap with each other.

We saw however that when a diatomic is stretched far beyond its equilibrium bond length, significant differences between the constrained/implicit OEP potentials are indeed observed, with step-like features appearing in the vicinity of the atom where the xc screening hole is not localised. These steps are a feature of the exact KS potential which are hard to capture using (semi-)local functionals, so their appear-

ance here is both expected and very promising. Although it was observed that these steps are too small to completely correct for the fractional charge dissociation error, this method provides progress towards systematically combating this error in commonly used functionals.

Chapter 8

Conclusions and further work

8.1 Conclusions

At the core of the work presented in this thesis is the constrained OEP method, introduced and improved upon in § 4. This method was shown to systematically improve calculated properties from (semi-)local density functional approximations by correcting the asymptotic decay of the xc potential. The newly implemented version of the constrained OEP code was shown to improve efficiency by streamlining the routines and lifting the positivity constraint. The ability to relax this constraint enforced on the OEP potential in the original implementations of the constrained method was explored in detail and shown to be possible for systems with more than a few electrons and through the use of modest-sized auxiliary basis sets, removing a major computational bottleneck on these calculations.

In § 5, we discussed the historical disconnect between KS DFT and spin-DFT in the limit of zero applied magnetic field. By deriving a new set of GKS equations we addressed this issue and connected the two exact theories in this limit, in the process deriving the exact DFT functional for the observable spin-density. As a consequence, a systematic error present in unpolarised DFAs when describing open-shell systems was identified, which was termed the ghost exchange error. A new approach was developed for constructing new DFAs which do not suffer

from this error and implement this OEP method into the HIPPO finite basis set code. Independent studies were then undertaken to understand the improvements in energies and system properties obtained using these new DFAs. Ground state total energies, ionisation energies, and electron affinities for open-shell systems all exhibited systematic accuracy gains relative to the corresponding unpolarised DFA. It was also seen that accurate non-OEP approximations to this theory can be made, bypassing the computational cost invoked by solving the OEP equation.

Using work from the two previous OEP formulations, in § 6 we derived an extension to the OEP scheme for open-shells to systems with a fractional number of electrons. This new ensemble OEP scheme was shown to yield ensemble energies that varied linearly with respect to electron number and to correctly exhibit a derivative discontinuity in the ensemble energy at integer values of electrons for (semi-)local DFAs. We also explored the variation in the ensemble potentials calculated where an infinitesimal amount of an electron is both added and removed; an alternative way to estimate the xc derivative discontinuity. It was observed that the ensemble potential varies by a constant in the vicinity of the system, which decays at large distances and converges well with respect to the size of the auxiliary basis set. This shift occurs as a direct consequence of the discontinuity in the effective screening charge of the ensemble system, which increases from $N-1$ to N as the number of electrons surpasses N.

Finally, in § 7 a novel way to reformulate the constrained method was introduced, by separating the total effective screening density of a system into separate screening densities localised on individual atoms. Each of these atomically localised screening densities is expanding in terms of the auxiliary basis functions localised on the specific atom, i.e. in a subset of the total set of auxiliary basis functions. For small separations, this method was shown to reduce to the original constrained method as expected. However, OEP potentials calculated when the xc screening hole was localised onto the atom with the smallest electron affinity of a heteronuclear diatomic stretched far from its equilibrium bond length (where the atom-

ically localised auxiliary basis functions do not significantly overlap), were seen to exhibit constant shifts in the neighbourhood of the atom with the higher affinity.

8.2 Future work

8.2.1 Modelling charge-transfer excitations

The method presented in § 7 provides a novel new way to solve the OEP equation with effective screening charges localised on individual atoms by decomposing the total molecular screening charge. As we saw, results obtained from this method produced some very exciting results, with the emergence of exotic step-like structures in the xc potential being of particular interest.

One potential application of this method is for the study of charge-transfer excitations, where an electron moves from one part of a system to another. These excitations are hard to study with standard approaches in DFT using (semi-)local functionals, as the inter-atomic steps in the xc potential are expected to be necessary to drive this kind of transfer of charge.

When studying these systems, we must consider individual parts of the molecule as separate entities, and so ability to the decompose the total molecular charge into pieces localised to different parts of system opens up other potential applications.

Due to the time restrictions of this study, it was not possible to undertake a study of this application, though it is nevertheless an exciting route for further research.

8.2.2 Understanding the decay of the derivative discontinuity

An additional clear extension to the work presented in this thesis relates to the work on estimating the derivative discontinuity via the difference between the ensemble potentials for the $N + \omega$ and $N - \omega$, $\omega \to 0$, systems.

In § 6, we observed that Δ_{xc} calculated from these potential differences manifests as a constant shift only within the vicinity of the system, decaying to zero as $1/r$ as $r \to \infty$. The spatial extent of this shift has not yet been fully addressed and a more work is required to understand this fully.

Originally, it was thought that this decay was a result of deficiencies in the underlying (semi-)local density functional approximations employed, however, as mentioned in § 6.2.3, a separate study on the inversion of full configuration interaction ensemble densities highlighted the same feature; a localised shift that decays outside the neighbourhood of the system. Therefore, it was concluded that this is in fact a limitation of a localised Gaussian basis set implementation.

To allow the shift to extend beyond the region spanned by the system, we could include basis functions localised away from the system, localised on 'ghost' atoms (fictitious atoms with no nuclei or electrons), that surround the system in a spherical arrangement. This would therefore allow for the flexibility to represent the derivative discontinuity away from the system and to extend it beyond the $1/r$ decay line which limits its extent in the current implementation.

8.3 Final Remarks

Throughout this thesis several important advancements in our understanding of DFT and its approximations through the development of new theoretical frameworks and OEP schemes were introduced.

Perhaps most importantly, we determined the first exact functional of an observable quantity in DFT, the spin-density, aside from the density and total energy, and in the process provided a link between in the exact theories of DFT and spin-DFT that had remained unknown for decades.

We also highlighted how extensions to the constrained OEP method to ensemble systems can yield non-zero derivative discontinuities using (semi-)local functionals,

and introduced a novel method for decomposing the molecular screening charge, resulting in the emergence of inter-atomic steps in the xc potential. Together, these two new OEP schemes provide important insights and exciting progress that will be applied to a wide range of systems in the future.

Bibliography

[1] J. Centrella, J. G. Baker, B. J. Kelly and J. R. van Meter, "Black-hole binaries, gravitational waves, and numerical relativity", Rev. Mod. Phys. **82**, 3069–3119 (2010).

[2] *Obr macroeconomic model*, `https://obr.uk/forecasts-in-depth/obr-macroeconomic-model/`, Accessed: 2023-01-04.

[3] H. Li, T. Ilyina, T. Loughran, A. Spring and J. Pongratz, "Reconstructions and predictions of the global carbon budget with an emission-driven earth system model", Earth System Dynamics Discussions **2022**, 1–26 (2022).

[4] N. Ferguson, D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. Van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, C. Walters, P. Winskill, C. Whittaker, C. Donnelly, S. Riley and A. Ghani, "Report 9: impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand", 1–20 (2020).

[5] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas", Phys. Rev. **136**, B864–B871 (1964).

[6] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects", Phys. Rev. **140**, A1133–A1138 (1965).

[7]R. Haunschild, A. Barth and B. French, "A comprehensive analysis of the history of dft based on the bibliometric method rpys", Journal of Cheminformatics **11**, 72 (2019).

[8]N. Mardirossian and M. Head-Gordon, "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals", Molecular Physics **115**, 2315–2372 (2017).

[9]P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields", The Journal of Physical Chemistry **98**, 11623–11627 (1994).

[10]J. P. Perdew, K. Burke and M. Ernzerhof, "Generalized gradient approximation made simple", Phys. Rev. Lett. **77**, 3865–3868 (1996).

[11]R. M. Martin, *Electronic structure: basic theory and practical methods* (Cambridge University Press, 2008).

[12]M. Born and R. Oppenheimer, "Zur quantentheorie der molekeln", Annalen der Physik **389**, 457–484 (1927).

[13]S. Matsika and P. Krause, "Nonadiabatic events and conical intersections", Annual Review of Physical Chemistry **62**, PMID: 21219147, 621–643 (2011).

[14]N. I. Gidopoulos and E. K. U. Gross, "Electronic non-adiabatic states: towards a density functional theory beyond the born-oppenheimer approximation", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372**, 20130059 (2014).

[15]D. R. Hartree, "The wave mechanics of an atom with a non-coulomb central field. part i. theory and methods", Mathematical Proceedings of the Cambridge Philosophical Society **24**, 89–110 (1928).

[16]D. R. Hartree, "The wave mechanics of an atom with a non-coulomb central field. part ii. some results and discussion", Mathematical Proceedings of the Cambridge Philosophical Society **24**, 111–132 (1928).

[17]D. R. Hartree, "The wave mechanics of an atom with a non-coulomb central field. part iii. term values and intensities in series in optical spectra", Mathematical Proceedings of the Cambridge Philosophical Society **24**, 426–437 (1928).

[18]J. C. Slater, "The theory of complex spectra", Phys. Rev. **34**, 1293–1322 (1929).

[19]E. U. Condon, "The theory of complex spectra", Phys. Rev. **36**, 1121–1133 (1930).

[20]V. Fock, "Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems", Zeitschrift fur Physik **61**, 126–148 (1930).

[21]P. Echenique and J. L. Alonso, "A mathematical and computational review of hartree–fock scf methods in quantum chemistry", Molecular Physics **105**, 3057–3098 (2007).

[22]C. C. J. Roothaan, "New developments in molecular orbital theory", Rev. Mod. Phys. **23**, 69–89 (1951).

[23]"The molecular orbital theory of chemical valency viii. a method of calculating ionization potentials", Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **205**, 541–552 (1951).

[24]T. Koopmans, "Über die zuordnung von wellenfunktionen und eigenwerten zu den einzelnen elektronen eines atoms", Physica **1**, 104–113 (1934).

[25]P. C. Coulson and M. I. Fischer, "Xxxiv. notes on the molecular orbital treatment of the hydrogen molecule", The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **40**, 386–393 (1949).

[26]C. C. J. Roothaan, "Self-consistent field theory for open shells of electronic systems", Rev. Mod. Phys. **32**, 179–185 (1960).

[27]H. B. Schlegel, "Potential energy curves using unrestricted mo/ller–plesset perturbation theory with spin annihilation", The Journal of Chemical Physics **84**, 4530–4534 (1986).

[28]J. S. Andrews, D. Jayatilaka, R. G. Bone, N. C. Handy and R. D. Amos, "Spin contamination in single-determinant wavefunctions", Chemical Physics Letters **183**, 423–431 (1991).

[29]P.-O. Löwdin, J. R. Sabin, M. C. Zerner and E. Brändas, eds., *The configuration interaction method: advances in highly correlated approaches*, Vol. 34, Advances in Quantum Chemistry (Academic Press, 1999), pp. 143–269.

[30]C. Møller and M. S. Plesset, "Note on an approximation treatment for many-electron systems", Phys. Rev. **46**, 618–622 (1934).

[31]J. Čıéžek, "On the use of the cluster expansion and the technique of diagrams in calculations of correlation effects in atoms and molecules", in *Advances in chemical physics* (John Wiley and Sons, Ltd, 1969), pp. 35–89.

[32]M. P. Nightingale and C. J. Umrigar, *Quantum monte carlo methods in physics and chemistry* (Springer, 1999).

[33]L. H. Thomas, "The calculation of atomic fields", Mathematical Proceedings of the Cambridge Philosophical Society **23**, 542–548 (1927).

[34]E. Fermi, "Un metodo statistico per la determinazione di alcune priorieta dell'atome", Rend. Accad. Naz. Lincei **6**, 32 (1927).

[35]B. Zhou, V. L. Ligneres and E. A. Carter, "Improving the orbital-free density functional theory description of covalent materials", The Journal of Chemical Physics **122**, 044103 (2005).

[36]M. Levy, "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem", Proceedings of the National Academy of Sciences **76**, 6062–6065 (1979).

[37]M. Levy, "Electron densities in search of hamiltonians", Phys. Rev. A **26**, 1200–1208 (1982).

[38]E. H. Lieb, "Density functionals for coulomb systems", International Journal of Quantum Chemistry **24**, 243–277 (1983).

[39]J. E. Harriman, "Orthonormal orbitals for the representation of an arbitrary density", Phys. Rev. A **24**, 680–682 (1981).

[40]R. G. Parr, R. A. Donnelly, M. Levy and W. E. Palke, "Electronegativity: the density functional viewpoint", The Journal of Chemical Physics **68**, 3801–3807 (1978).

[41]N. Ashcroft and N. Mermin, *Solid state physics* (Holt-Saunders, 1976).

[42]S. H. Vosko, L. Wilk and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis", Canadian Journal of Physics **58**, 1200–1211 (1980).

[43]J. P. Perdew and A. Zunger, "Self-interaction correction to density-functional approximations for many-electron systems", Phys. Rev. B **23**, 5048–5079 (1981).

[44]J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy", Phys. Rev. B **45**, 13244–13249 (1992).

[45]J. P. Perdew, "Accurate density functional for the energy: real-space cutoff of the gradient expansion for the exchange hole", Phys. Rev. Lett. **55**, 1665–1668 (1985).

[46]D. C. Langreth and J. P. Perdew, "Theory of nonuniform electronic systems. i. analysis of the gradient approximation and a generalization that works", Phys. Rev. B **21**, 5469–5493 (1980).

[47]D. C. Langreth and M. J. Mehl, "Beyond the local-density approximation in calculations of ground-state electronic properties", Phys. Rev. B **28**, 1809–1834 (1983).

[48]J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, "Atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation", Phys. Rev. B **46**, 6671–6687 (1992).

[49]A. D. Becke, "Density functional calculations of molecular bond energies", The Journal of Chemical Physics **84**, 4524–4529 (1986).

[50]A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior", Phys. Rev. A **38**, 3098–3100 (1988).

[51]C. Lee, W. Yang and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density", Phys. Rev. B **37**, 785–789 (1988).

[52]A. D. Becke, "Density-functional thermochemistry. i. the effect of the exchange-only gradient correction", The Journal of Chemical Physics **96**, 2155–2160 (1992).

[53]S. K. Ghosh and R. G. Parr, "Phase-space approach to the exchange-energy functional of density-functional theory", Phys. Rev. A **34**, 785–791 (1986).

[54]A. D. Becke and M. R. Roussel, "Exchange holes in inhomogeneous systems: a coordinate-space model", Phys. Rev. A **39**, 3761–3767 (1989).

[55]J. Sun, A. Ruzsinszky and J. P. Perdew, "Strongly constrained and appropriately normed semilocal density functional", Phys. Rev. Lett. **115**, 036402 (2015).

[56]A. D. Becke, "A new mixing of hartree–fock and local density-functional theories", The Journal of Chemical Physics **98**, 1372–1377 (1993).

[57]J. P. Perdew, M. Ernzerhof and K. Burke, "Rationale for mixing exact exchange with density functional approximations", The Journal of Chemical Physics **105**, 9982–9985 (1996).

[58]A. D. Becke, "Density-functional thermochemistry. iii. the role of exact exchange", The Journal of Chemical Physics **98**, 5648–5652 (1993).

[59]K. Burke, M. Ernzerhof and J. P. Perdew, "The adiabatic connection method: a non-empirical hybrid", Chemical Physics Letters **265**, 115–120 (1997).

[60]J. P. Perdew, W. Yang, K. Burke, Z. Yang, E. K. U. Gross, M. Scheffler, G. E. Scuseria, T. M. Henderson, I. Y. Zhang, A. Ruzsinszky, H. Peng, J. Sun, E. Trushin and A. Görling, "Understanding band gaps of solids in generalized kohn–sham theory", Proceedings of the National Academy of Sciences **114**, 2801–2806 (2017).

[61] A. Seidl, A. Görling, P. Vogl, J. A. Majewski and M. Levy, "Generalized kohn-sham schemes and the band-gap problem", Phys. Rev. B **53**, 3764–3774 (1996).

[62] U. von Barth and L. Hedin, "A local exchange-correlation potential for the spin polarized case. i", Journal of Physics C: Solid State Physics **5**, 1629–1642 (1972).

[63] K. Capelle and G. Vignale, "Nonuniqueness of the potentials of spin-density-functional theory", Phys. Rev. Lett. **86**, 5546–5549 (2001).

[64] H. Eschrig and W. Pickett, "Density functional theory of magnetic systems revisited", Solid State Communications **118**, 123–127 (2001).

[65] C. A. Ullrich, "Nonuniqueness in spin-density-functional theory on lattices", Phys. Rev. B **72**, 073102 (2005).

[66] N. I. Gidopoulos, "Potential in spin-density-functional theory of noncollinear magnetism determined by the many-electron ground state", Phys. Rev. B **75**, 134408 (2007).

[67] J. P. Perdew and K. Schmidt, "Jacob's ladder of density functional approximations for the exchange-correlation energy", in Density functional theory and its application to materials, Vol. 577, American Institute of Physics Conference Series (July 2001), pp. 1–20.

[68] R. T. Sharp and G. K. Horton, "A variational approach to the unipotential many-electron problem", Phys. Rev. **90**, 317–317 (1953).

[69] A. Görling and M. Levy, "Exact kohn-sham scheme based on perturbation theory", Phys. Rev. A **50**, 196–204 (1994).

[70] A. Heßelmann, A. W. Götz, F. Della Sala and A. Görling, "Numerically stable optimized effective potential method with balanced gaussian basis sets", The Journal of Chemical Physics **127**, 054102 (2007).

[71] A. Görling, A. Heßelmann, M. Jones and M. Levy, "Relation between exchange-only optimized potential and kohn–sham methods with finite basis sets, and effect of linearly dependent products of orbital basis functions", The Journal of Chemical Physics **128**, 104104 (2008).

[72]J. B. Krieger, Y. Li and G. J. Iafrate, "Construction and application of an accurate local spin-polarized kohn-sham potential with integer discontinuity: exchange-only theory", Phys. Rev. A **45**, 101–126 (1992).

[73]J. B. Krieger, Y. Li and G. J. Iafrate, "Systematic approximations to the optimized effective potential: application to orbital-density-functional theory", Phys. Rev. A **46**, 5453–5458 (1992).

[74]O. V. Gritsenko and E. J. Baerends, "Orbital structure of the kohn-sham exchange potential and exchange kernel and the field-counteracting potential for molecules in an electric field", Phys. Rev. A **64**, 042506 (2001).

[75]F. Della Sala and A. Görling, "Efficient localized hartree–fock methods as effective exact-exchange kohn–sham methods for molecules", The Journal of Chemical Physics **115**, 5718–5732 (2001).

[76]A. Unsöld, "Quantentheorie des wasserstoffmolekülions und der born-landéschen abstoßungskräfte", Zeitschrift für Physik **43**, 563–574 (1927).

[77]J. C. Slater, "Atomic shielding constants", Phys. Rev. **36**, 57–64 (1930).

[78]T. Kato, "On the eigenfunctions of many-particle systems in quantum mechanics", Communications on Pure and Applied Mathematics **10**, 151–177 (1957).

[79]N. H. March, "Spatially dependent generalization of kato's theorem for atomic closed shells in a bare coulomb field", Phys. Rev. A **33**, 88–89 (1986).

[80]J. M. Foster and S. F. Boys, "A quantum variational calculation for hcho", Rev. Mod. Phys. **32**, 303–304 (1960).

[81]W. J. Hehre, R. F. Stewart and J. A. Pople, "Self-consistent molecular-orbital methods. i. use of gaussian expansions of slater-type atomic orbitals", The Journal of Chemical Physics **51**, 2657–2664 (1969).

[82]C. M. Reeves and R. Fletcher, "Use of gaussian functions in the calculation of wavefunctions for small molecules. iii. the orbital basis and its effect on valence", The Journal of Chemical Physics **42**, 4073–4081 (1965).

[83]W. J. Hehre, R. F. Stewart and J. A. Pople, "Self-consistent molecular-orbital methods. i. use of gaussian expansions of slater-type atomic orbitals", The Journal of Chemical Physics **51**, 2657–2664 (1969).

[84]W. J. Hehre, R. Ditchfield and J. A. Pople, "Self—consistent molecular orbital methods. xii. further extensions of gaussian—type basis sets for use in molecular orbital studies of organic molecules", The Journal of Chemical Physics **56**, 2257–2261 (1972).

[85]T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen", The Journal of Chemical Physics **90**, 1007–1023 (1989).

[86]D. E. Woon and T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. iii. the atoms aluminum through argon", The Journal of Chemical Physics **98**, 1358–1371 (1993).

[87]A. D. Becke, "A multicenter numerical integration scheme for polyatomic molecules", The Journal of Chemical Physics **88**, 2547–2553 (1988).

[88]P. Pulay, "Convergence acceleration of iterative sequences. the case of scf iteration", Chemical Physics Letters **73**, 393–398 (1980).

[89]P. Pulay, "Improved scf convergence acceleration", Journal of Computational Chemistry **3**, 556–560 (1982).

[90]A. T. B. Gilbert, N. A. Besley and P. M. W. Gill, "Self-consistent field calculations of excited states using the maximum overlap method (mom)", The Journal of Physical Chemistry A **112**, 13164–13171 (2008).

[91]N. N. Lathiotakis and M. A. L. Marques, "Benchmark calculations for reduced density-matrix functional theory", The Journal of Chemical Physics **128**, 184103 (2008).

[92]M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis

and J. A. Montgomery Jr, "General atomic and molecular electronic structure system", Journal of Computational Chemistry **14**, 1347–1363 (1993).

[93]M. S. Gordon and M. W. Schmidt, "Chapter 41 - advances in electronic structure theory: gamess a decade later", in *Theory and applications of computational chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim and G. E. Scuseria (Elsevier, Amsterdam, 2005), pp. 1167–1189.

[94]M. A. L. Marques, M. J. T. Oliveira and T. Burnus, "Libxc: a library of exchange and correlation functionals for density functional theory", Comput. Phys. Commun. **183**, 2272–2281 (2012).

[95]B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson and T. L. Windus, "New basis set exchange: an open, up-to-date resource for the molecular sciences community", Journal of Chemical Information and Modeling **59**, 4814–4820 (2019).

[96]D. Feller, "The role of databases in support of computational chemistry calculations", Journal of Computational Chemistry **17**, 1571–1586 (1996).

[97]K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li and T. L. Windus, "Basis set exchange: a community database for computational sciences", Journal of Chemical Information and Modeling **47**, 1045–1052 (2007).

[98]R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, J. F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford and C. D. Sherrill, "Psi4 1.1: an open-source electronic structure program emphasizing automation, advanced libraries, and interoperability", Journal of Chemical Theory and Computation **13**, 3185–3197 (2017).

[99]U. Bozkaya and C. D. Sherrill, "Analytic energy gradients for the coupled-cluster singles and doubles with perturbative triples method with the density-fitting approximation", The Journal of Chemical Physics **147**, 044104 (2017).

[100]N. I. Gidopoulos and N. N. Lathiotakis, "Constraining density functional approximations to yield self-interaction free potentials", The Journal of Chemical Physics **136**, 224109 (2012).

[101]N. Gidopoulos and N. N. Lathiotakis, *Chapter 6 - constrained local potentials for self-interaction correction*, edited by E. Arimondo, C. C. Lin and S. F. Yelin, Vol. 64, Advances In Atomic, Molecular, and Optical Physics (Academic Press, 2015), pp. 129–142.

[102]T. J. Callow, B. J. Pearce, T. Pitts, N. N. Lathiotakis, M. J. P. Hodgson and N. I. Gidopoulos, "Improving the exchange and correlation potential in density-functional approximations through constraints", Faraday Discuss. **224**, 126–144 (2020).

[103]A. Görling, "New ks method for molecules based on an exchange charge density generating the exact local ks exchange potential", Phys. Rev. Lett. **83**, 5459–5462 (1999).

[104]J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz, "Density-functional theory for fractional particle number: derivative discontinuities of the energy", Phys. Rev. Lett. **49**, 1691–1694 (1982).

[105]J. P. Perdew and M. Levy, "Comment on "significance of the highest occupied kohn-sham eigenvalue"", Phys. Rev. B **56**, 16021–16028 (1997).

[106]R. van Meer, O. V. Gritsenko and E. J. Baerends, "Physical meaning of virtual kohn–sham orbitals and orbital energies: an ideal basis for the description of molecular excitations", Journal of Chemical Theory and Computation **10**, PMID: 26588140, 4432–4441 (2014).

[107]P. Verma and R. J. Bartlett, "Increasing the applicability of density functional theory. ii. correlation potentials from the random phase approximation and beyond", The Journal of Chemical Physics **136**, 044105 (2012).

[108]R. J. Bartlett, "Adventures in dft by a wavefunction theorist", The Journal of Chemical Physics **151**, 160901 (2019).

[109]M. Levy, J. P. Perdew and V. Sahni, "Exact differential equation for the density and ionization energy of a many-particle system", Phys. Rev. A **30**, 2745–2748 (1984).

[110]C.-O. Almbladh and U. von Barth, "Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues", Phys. Rev. B **31**, 3231–3244 (1985).

[111]R. M. Dreizler and E. K. U. Gross, *Density functional theory: an approach to the quantum many-body problem* (Springer, Berlin, 1990).

[112]T. Pitts, N. I. Gidopoulos and N. N. Lathiotakis, "Performance of the constrained minimization of the total energy in density functional approximations: the electron repulsion density and potential", The European Physical Journal B **91**, 130 (2018).

[113]N. I. Gidopoulos and N. N. Lathiotakis, "Nonanalyticity of the optimized effective potential with finite basis sets", Phys. Rev. A **85**, 052508 (2012).

[114]R. Johnson, "Nist 101. computational chemistry comparison and benchmark database", Accessed: 2023-01-10 (1999).

[115]T. J. Callow, B. Pearce and N. I. Gidopoulos, "Density functionals with spin-density accuracy for open shells", The Journal of Chemical Physics **156**, 111101 (2022).

[116]R. G. Parr and R. G. Yang, *Density functional theory of atoms and molecules* (Oxford University Press, 1989) Chap. 8, pp. 173–174.

[117]T. J. Callow, "Systematic routes to improved approximations in kohn–sham theory", Available at Durham E-Theses Online: http://etheses.dur.ac.uk/13817 (2020).

[118]E. Runge and E. K. U. Gross, "Density-functional theory for time-dependent systems", Phys. Rev. Lett. **52**, 997–1000 (1984).

[119]M. Casida and M. Huix-Rotllant, "Progress in time-dependent density-functional theory", Annual Review of Physical Chemistry **63**, PMID: 22242728, 287–323 (2012).

[120]L. Schimka, J. Harl and G. Kresse, "Improved hybrid functional for solids: the hsesol functional", The Journal of Chemical Physics **134**, 024116 (2011).

[121]F. Tran and P. Blaha, "Accurate band gaps of semiconductors and insulators with a semilocal exchange-correlation potential", Phys. Rev. Lett. **102**, 226401 (2009).

[122]T. Stein, H. Eisenberg, L. Kronik and R. Baer, "Fundamental gaps in finite systems from eigenvalues of a generalized kohn-sham method", Phys. Rev. Lett. **105**, 266802 (2010).

[123]G. Onida, L. Reining and A. Rubio, "Electronic excitations: density-functional versus many-body green's-function approaches", Rev. Mod. Phys. **74**, 601–659 (2002).

[124]E. K. U. Gross, L. N. Oliveira and W. Kohn, "Rayleigh-ritz variational principle for ensembles of fractionally occupied states", Phys. Rev. A **37**, 2805–2808 (1988).

[125]E. K. U. Gross, L. N. Oliveira and W. Kohn, "Density-functional theory for ensembles of fractionally occupied states. i. basic formalism", Phys. Rev. A **37**, 2809–2820 (1988).

[126]B. Senjean and E. Fromager, "N-centered ensemble density-functional theory for open systems", International Journal of Quantum Chemistry **120**, e26190 (2020).

[127]C. Marut, B. Senjean, E. Fromager and P.-F. Loos, "Weight dependence of local exchange–correlation functionals in ensemble density-functional theory: double excitations in two-electron systems", Faraday Discuss. **224**, 402–423 (2020).

[128]F. Cernatic, B. Senjean, V. Robert and E. Fromager, "Ensemble density functional theory of neutral and charged excitations", Topics in Current Chemistry **380**, 4 (2021).

[129]J. F. Janak, "Proof that de/dn=e in density-functional theory", Phys. Rev. B **18**, 7165–7168 (1978).

[130]S. Bousiadi and N. Nektarios, to be submitted.

[131]A. J. Cohen, P. Mori-Sánchez and W. Yang, "Insights into current limitations of density functional theory", Science **321**, 792–794 (2008).

[132]P. Mori-Sánchez, A. J. Cohen and W. Yang, "Localization and delocalization errors in density functional theory and implications for band-gap prediction", Phys. Rev. Lett. **100**, 146401 (2008).

[133]M. J. P. Hodgson, J. D. Ramsden and R. W. Godby, "Origin of static and dynamic steps in exact kohn-sham potentials", Phys. Rev. B **93**, 155146 (2016).

[134]M. J. P. Hodgson, E. Kraisler, A. Schild and E. K. U. Gross, "How interatomic steps in the exact kohn–sham potential relate to derivative discontinuities of the energy", The Journal of Physical Chemistry Letters **8**, 5974–5980 (2017).

[135]H. Hotop and W. C. Lineberger, "Binding energies in atomic negative ions: ii", Journal of Physical and Chemical Reference Data **14**, 731–750 (1985).

[136]U. Berzinsh, M. Gustafsson, D. Hanstorp, A. Klinkmüller, U. Ljungblad and A.-M. Mårtensson-Pendrill, "Isotope shift in the electron affinity of chlorine", Phys. Rev. A **51**, 231–238 (1995).

[137]K. R. Lykke, K. K. Murray and W. C. Lineberger, "Threshold photodetachment of H^-", Phys. Rev. A **43**, 6104–6107 (1991).

[138]G. Haeffler, D. Hanstorp, I. Kiyan, A. E. Klinkmüller, U. Ljungblad and D. J. Pegg, "Electron affinity of li: a state-selective measurement", Phys. Rev. A **53**, 4127–4131 (1996).

[139]R. S. Mulliken, "Electronic population analysis on lcao–mo molecular wave functions. ii. overlap populations, bond orders, and covalent bond energies", The Journal of Chemical Physics **23**, 1841–1846 (1955).